

126578

From: Borin, Michael  
Sent: Wednesday, July 07, 2004 11:51 AM  
To: STIC-Biotech/ChemLib  
Subject: Search request: 10/081108

CRFG

Examiner: M.Borin  
AU: 1631  
Mailbox:2C70  
Office: Remsen 2A55  
Tel.: 20713

RE: 10/081108

Please search :

1. nucleic acids of SEQ ID NO: 1.
2. polypeptide SEQ ID NO: 2.

against the commercial protein and nucleic acid databases + interference search.

Thank you

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: *July 7, 2004*  
Date Completed: *July 7, 2004*  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_ /  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: *olyp kosp*  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**This Page Blank (uspto)**



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 126578

**TO:** Michael Borin  
**Location:** REM-2A55&2C70  
**Art Unit:** 1631  
**Friday, July 09, 2004**

**Case Serial Number:** 10/081108

**From:** Edward Hart  
**Location:** Biotech-Chem Library  
**REM-1A55**  
**Phone:** 571-272-2512

[edward.hart@uspto.gov](mailto:edward.hart@uspto.gov)

### Search Notes

Examiner Borin,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**This Page Blank (uspto)**

Page 6







DE Hypothetical protein.  
 OS Homo sapiens (Human). Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Bivalvia; Mollusca; Ciliophora; Rhizarians; Amoebozoa; Radiolaria; Forams;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Strausberg R.; Submitted (May-2001) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; BC000404; AAH08404.1; -.  
 DR GO; GO:0005829; Cytosol; ISS.  
 GO; GO:0005624; C:membrane fraction; ISS.  
 DR GO; GO:001050; F:D-erythro-sphingosine kinase activity; ISS.  
 DR GO; GO:0000287; F:magnesium ion binding; ISS.  
 DR GO; GO:0007242; P:intracellular signaling cascade; ISS.  
 DR GO; GO:0046521; P:sphingoid catabolism; ISS.  
 DR InterPro; IPR001395; Aldo/ket\_red.  
 DR InterPro; IPR001206; DAGKC.  
 Pfam; PF00781; DAGKC; 1.  
 ProDom; PD005043; DAGKC; 1.  
 SMART; SM00046; DAGKC; 1.  
 PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.  
 KW Hypothetical Protein.  
 SQ SEQUENCE 384 AA; 42474 MW; F52999FF306113B0 CRC64;

Query Match 29.1%; Score 62; DB 4; Length 384;  
 Best Local Similarity 40.0%; Pred. No. 3.1;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLAASQQLQARLKKES-----PVVSMWLEPDG 36  
 Db 295 SRAMLLRFLAMERGRHMEYECPLVTVPPVAFRLPEPDG 334

RESULT 3  
 Q96GKL ID Q96GKL PRELIMINARY; PRT; 398 AA.  
 AC Q96GKL; TRMBlrel. 19, Created  
 DT 01-DEC-2001 (TRMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TRMBlrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.; Submitted (Jun-2001) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; BC009419; AAH09419.1; -.  
 DR GO; GO:0005829; Cytosol; ISS.  
 DR GO; GO:0005624; C:membrane fraction; ISS.  
 DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; ISS.  
 DR GO; GO:0007242; P:intracellular signaling cascade; ISS.  
 DR GO; GO:0046521; P:sphingoid catabolism; ISS.  
 DR InterPro; IPR001395; Aldo/ket\_red.  
 DR InterPro; IPR01216; DAGKC.  
 Pfam; PF00781; DAGKC; 1.  
 ProDom; PD005043; DAGKC; 1.  
 SMART; SM00046; DAGKC; 1.  
 PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.

RESULT 4  
 Q9N632 ID Q9N632 PRELIMINARY; PRT; 470 AA.  
 AC Q9N632; TRMBlrel. 22, Created  
 DT 01-OCT-2002 (TRMBlrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TRMBlrel. 25, Last annotation update)  
 DE Similar to sphingosine kinase.  
 OS Homo sapiens (Human).  
 OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood, and Skin;  
 RA Strausberg R.; Submitted (May-2002) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; BC0010553; AAH30553.1; -.  
 DR EMBL; BC014429; AAH14429.1; -.  
 DR GO; GO:0005829; C:cytosol; ISS.  
 DR GO; GO:0005624; C:membrane fraction; ISS.  
 DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; ISS.  
 DR GO; GO:000287; F:magnesium ion binding; ISS.  
 DR GO; GO:0046521; P:intracellular signaling cascade; ISS.  
 DR InterPro; IPR001395; Aldo/ket\_red.  
 DR InterPro; IPR001206; DAGKC.  
 Pfam; PF00781; DAGKC; 1.  
 ProDom; PD005043; DAGKC; 1.  
 SMART; SM00046; DAGKC; 1.  
 PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.  
 KW Kinase.  
 SQ SEQUENCE 470 AA; 51064 MW; 5172E93A38C7/C17 CRC64;

Query Match 29.1%; Score 62; DB 4; Length 470;  
 Best Local Similarity 40.0%; Pred. No. 3.8;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLAASQQLQARLKKES-----PVVSMWLEPDG 36  
 Db 381 SRAMLLRFLAMERGRHMEYECPLVTVPPVAFRLPEPDG 420

RESULT 5  
 Q87PB1 ID Q87PB1 PRELIMINARY; PRT; 668 AA.  
 AC Q87PB1; TRMBlrel. 24, Created  
 DT 01-JUN-2003 (TRMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TRMBlrel. 25, Last annotation update)  
 DE Tail-specific protease.  
 CN VP1606.  
 OS Vibrio parahaemolyticus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=670;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RIMD 210633 / Serotype O3:K6;  
 RX MEDLINE=22508154; PubMed=1262073;  
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
 RA Iijima Y., Naito M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
 RA Yasunaga T., Honda T., Shinsugawa H., Hattori M., Iida T.;  
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism  
 RL distinct from that of V. cholerae.";  
 DR AP005078; BAD5986.1; -.  
 DR GO; GO:0008236; Peptidase activity; IEA.  
 DR GO; GO:0007242; Peptidase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model 1

Run on: July 7, 2004, 17:04:24 ; Search time 39 Seconds  
(without alignments)  
347.879 Million cell updates/sec

Title: US-10-081-108-2  
Perfect score: 213  
Sequence: I MAARAYFLAQLSAQLLQARLM.....SPVYSWRLPEPDGTALCFIF 43

Scoring table: BLOSUM62  
Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:  
 1: sp\_archaea:  
 2: sp\_bacteria:  
 3: sp\_fungi:  
 4: sp\_human:  
 5: sp\_invertebrate:  
 6: sp\_mammal:  
 7: sp\_mhc:  
 8: sp\_organelle:  
 9: sp\_phage:  
 10: sp\_plant:  
 11: sp\_rabbit:  
 12: sp\_virus:  
 13: sp\_vertebrate:  
 14: sp\_unclassified:  
 15: sp\_virus:  
 16: sp\_bacteriap:  
 17: sp\_archeap:

SPTREMBL 25:  
 1: sp\_archaea:  
 2: sp\_bacteria:  
 3: sp\_fungi:  
 4: sp\_human:  
 5: sp\_invertebrate:  
 6: sp\_mammal:  
 7: sp\_mhc:  
 8: sp\_organelle:  
 9: sp\_phage:  
 10: sp\_plant:  
 11: sp\_rabbit:  
 12: sp\_virus:  
 13: sp\_vertebrate:  
 14: sp\_unclassified:  
 15: sp\_virus:  
 16: sp\_bacteriap:  
 17: sp\_archeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	62	29.1	290	4	Q9BTG7	Q9bg7 homo sapien
2	62	29.1	384	4	Q96HK8	Q96hv8 homo sapien
3	62	29.1	398	4	Q96GK1	Q96gk1 homo sapien
4	62	29.1	470	4	Q8N632	Q8n32 homo sapien
5	60	28.2	668	16	Q87PB1	Q8pb1 vibrio para
6	60	28.2	552	16	Q8ZQE6	Q8zqe6 salmonella
7	58	27.2	173	5	Q9VYCL	Q9vycl drosophila
8	57	27.0	619	16	Q89UX1	Q8aux1 bradyrhizob
9	57	26.8	340	16	Q7U9Q3	Q7q9q3 synechococcus
10	56	26.3	404	16	Q7URB8	Q7urb8 rhodopirell
11	56	26.3	2019	5	Q9W0D3	Q9wd3 drosophila
12	56	26.3	4899	5	Q9VRJ1	Q9vrj1 drosophila
13	55.5	26.1	126	11	Q9CRJ6	Q9crj6 mus musculus
14	55.5	26.1	317	16	Q82P22	Q82p2 streptomyces
15	55.5	26.1	359	11	Q9JMD5	Q9jmd5 mus musculus
16	55.5	26.1	359	11	Q9JH8	Q9jh8 mus musculus

## ALIGNMENTS

RESULT 1						
ID	Q9BTG7	PRELIMINARY;	PRT;	290 AA.		
AC	Q9BTG7;	PRELIMINARY;	PRT;	290 AA.		
DT	01-JUN-2001 (TREMBLrel. 1.7, Created)	04/05/01	04/05/01	04/05/01		
DT	01-JUN-2001 (TREMBLrel. 1.7, Last sequence update)	04/05/01	04/05/01	04/05/01		
DT	01-JUN-2003 (TREMBLrel. 2.4, Last annotation update)	04/05/03	04/05/03	04/05/03		
DE	Similar to sphingosine kinase 1 (Fragment).					
OS	Homo sapiens.					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
NCBI-TaxID	9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE-Skin;					
RA	Straubberg R.;					
RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBj databases.					
DR	EMBL; BC004112; AAH041121; -					
DR	GO; GO:0016301; F:Kinase activity; IEA.					
KW						
KW						
PT	NON-TER	1	1	1	1	
SQ	SEQUENCE	290 AA;	31998 MW;	A3BA219DB52CS20 CRC64;		
Query Match		29.1%	Score 62;	DB 4;	Length 290;	
Best Local Similarity		40.0%	Pre 0.23;			
Matches		16;	Conservative 8;	Mismatches 10;	Indels 6;	Gaps 1;
Qy	3 ARAVFIALSQAQLQARIMKEES---PVNSWRLEPEDG 36	:	:	:	:	:
Db	201 SRAMLLRFLAMEKGGRHEMYRCPYLYVVPVVAFRIEPKDG 240					

RESULT 2	Score	Query	Match	Length	DB ID	Description
		Q96HV8	PRELIMINARY;	PRT;	384 AA.	
		AC	Q96HV8;	PRELIMINARY;	01-DEC-2001 (TREMBLrel. 1.9, Created)	
		DT	01-DEC-2001 (TREMBLrel. 1.9, Last sequence update)	01-DEC-2001 (TREMBLrel. 1.9, Last sequence update)	01-OCT-2003 (TREMBLrel. 2.5, Last annotation update)	

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 23:35:12 ; Search time 6472 Seconds  
(without alignment)  
6723.796 Million cell updates/sec

Title: US-10-081-108-1  
Perfect score: 1004  
Sequence: 1 CGCCTAATTAGGTCTCCGG.....ATCTTTACACTAAAAAGCC 1004

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank:  
1: gb\_ba:\*

2: gb\_itg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_Dl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vii:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_ln:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_Ov:\*

22: em\_dat:\*

23: em\_Dh:\*

24: em\_Dl:\*

25: em\_ro:\*

26: em\_sits:\*

27: em\_in:\*

28: em\_vii:\*

29: em\_htg\_hum:\*

30: em\_htg\_inv:\*

31: em\_htg\_other:\*

32: em\_htg\_mus:\*

33: em\_htg\_Dln:\*

34: em\_htg\_rod:\*

35: em\_htg\_mam:\*

36: em\_htg\_rvt:\*

37: em\_sy:\*

38: em\_htgo\_hum:\*

39: em\_htgo\_mus:\*

40: em\_htgo\_other:\*

41: em\_htgo\_other:\*

Score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1004	100.0	1004	6	ARI107991		ARI107991 Sequence
2	1004	100.0	1004	6	ARI167384		ARI167384 Sequence
3	1004	100.0	1004	6	ARI236412		ARI236412 Sequence
4	1004	100.0	1004	6	BDI11457		BDI11457 Sequence
5	1004	100.0	1004	6	BDI176601		BDI176601 Method of
6	1004	100.0	1004	9	HSU1980		HSU1980 Human_B_mel
7	966	96.2	1032	6	ARI104146		ARI104146 Sequence
8	966	96.2	1032	6	I28470		I28470 Sequence 1
9	966	96.2	1032	6	I72215		I72215 Sequence 1
10	829.2	82.6	2154	9	AF527550		AF527550 Homo sapi
11	563.4	56.1	1068	9	AF527553		AF527553 Homo sapi
12	532.4	53.0	177896	2	AC064811		AC064811 Homo sapi
13	529.2	52.7	139063	2	AC133563		AC133563 Homo sapi
14	526	52.4	178300	2	AC145613		AC145613 Homo sapi
15	524.4	52.2	107701	2	AF155775		AF155775 Homo sapi
16	524.4	52.2	158712	2	AC074394		AC074394 Homo sapi
17	524.4	52.2	182078	2	AC140304		AC140304 Homo sapi
18	522.8	52.1	69684	9	AC006175		AC006175 Homo sapi
19	522.8	52.1	183396	2	AL161418		AL161418 Homo sapi
20	522.8	52.1	231556	2	AL158811		AL158811 Homo sapi
21	521.2	51.9	190277	9	AF254933		AF254933 Homo sapi
22	521.2	51.9	281116	9	HS21C001		HS21C001 Homo sapi
23	467.8	46.6	148606	9	AC104992		AC104992 Homo sapi
24	464.2	46.2	2267	9	AF521552		AF521552 Homo sapi
25	448	44.6	1104	9	AF527551		AF527551 Homo sapi
26	435.2	43.3	1840	9	AF339516		AF339516 Homo sapi
27	433.6	43.2	1840	9	AF339515		AF339515 Homo sapi
28	385	38.3	2071	9	AF527554		AF527554 Homo sapi
29	266	26.5	1891	9	AF339514		AF339514 Homo sapi
30	264.4	26.3	1891	9	AF18570		AF18570 Homo sapi
31	215.4	21.5	6318	9	AF499647		AF499647 Homo sapi
32	205.8	20.5	75657	9	HSB11JC8		HSB11JC8 Human DNA
33	205.8	20.5	182078	2	AC140904		AC140904 Homo sapi
34	199.4	19.9	41185	9	HSB1Lc16		HSB1Lc16
35	181.8	18.1	57728	6	E54646		E54646 Base sequen
36	181.8	18.1	129800	2	AL163539		AL163539 Homo sapi
37	181.8	18.1	14200	2	BX664725		BX664725 Homo sapi
38	181.8	18.1	8088717	1	BX088717		BX088717 Human DNA
39	181.8	18.1	202891	9	BX088551		BX088551 Human DNA
40	179.4	17.9	42108	9	AC138036		AC138036 Homo sapi
41	179.4	17.9	120515	9	AC008443		AC008443 Homo sapi
42	179.4	17.9	236913	2	AL161615		AL161615 Homo sapi
43	179.2	17.8	111648	9	HS16C2		HS16C2 Human DNA
44	179.2	17.8	190690	2	AL627334		AL627334 Homo sapi
45	177.6	17.7	119182	9	HSB62L20		HSB62L20 Homo sapi

## ALIGNMENTS

RESULT 1	AR107991	LOCUS	Sequence 1 from patent US 6110694.	DNA	Linear	PAT 14-FEB-2001
DEFINITION						
ACCESSION	AR107991	VERSION	AR107991.1	GI:12823478		
KEYWORDS						
SOURCE						
ORGANISM						
Unclassified						
REFERENCE	1	(bases 1 to 1004)				
AUTHORS	Boel,P., Wildmann,C., Boon-Falleur,T., van der Bruggen,P., Coulie,P. and Renaud,J.-C.					
TITLE	Methods for determining complexes of tumor rejection antigens and HLA-Cw*160/molecules					

Pred. No. is the number of results predicted by chance to have a

JOURNAL	Patent:	US 6110694-A-1 29-AUG-2000;
FEATURES	Location/Qualifiers	
source		
ORIGIN		
Query	Match	Score 1004; Score 1004; Score 1004;
	Best Local Similarity	100.0%; Pred. No. 3.7
	Matches 1004;	Conservative 0; Mismatches 1.
	source	.1004 /organism="unknown" /mol_type="unassigned DNA"
Qy	1 CGCCAAATTAGGCTTCTCGGTATTCGGCTGAG	
Db	1 CGCCAAATTAGGCTTCTCGGTATTCGGCTGAG	
Qy	61 AGGGAAAGGGGAACCTGAGGGCTGAGGCTGAA	
Db	61 AGGGAAAGGGGAACCTGAGGGCTGAGGCTGAA	
Qy	121 GTGGTGGCAACAGGATGCCAGGCAGTGGCAGTGAGT	
Db	121 GTGGTGGCAACAGGATGCCAGGCAGTGGCAGTGAGT	
Qy	181 AGTGGGGCTGGAGGAGTAAAGATCGGGCCAAGCC	
Db	181 AGTGGGGCTGGAGGAGTAAAGATCGGGCCAAGCC	
Qy	241 TGCCTCAAGCCAGGCTGTGTAAGGGAGGTCCCC	
Db	241 TGCCTCAAGCCAGGCTGTGTAAGGGAGGTCCCC	
Qy	301 AAGGGGCAAGCTGCTGCTTCACTCTCTGAGG	
Db	301 AAGGGGCAAGCTGCTGCTTCACTCTCTGAGG	
Qy	361 GCAGGTCAAGGGCAATAAGGGAGATGGACT	
Db	361 GCAGGTCAAGGGCAATAAGGGAGATGGACT	
Qy	421 GATCTCCGTGACCTGTGATGTCGCGCTTGGCC	
Db	421 GATCTCCGTGACCTGTGATGTCGCGCTTGGCC	
Qy	481 GTGGATTGTTAACCTTGAGGCCACTATCAA	
Db	481 GTGGATTGTTAACCTTGAGGCCACTATCAA	
Qy	541 TGTATCATTTATCCTGTGTGAGAGCGGGTC	
Db	541 TGTATCATTTATCCTGTGTGAGAGCGGGTC	
Qy	601 TGCTTGTGTCAGAACACATTGACAAAGTCTCTCA	
Db	601 TGCTTGTGTCAGAACACATTGACAAAGTCTCTCA	
Qy	661 TTAAACTTCACCAATTTAAGTCAAATAAACATA	
Db	661 TTAAACTTCACCAATTTAAGTCAAATAAACATA	
Qy	721 TATTCCTACTGTATTAAGGTAAATAACATAAA	
Db	721 TATTCCTACTGTATTAAGGTAAATAACATAAA	
Qy	781 AGAATTGTGTTGGAGGAAATGTATTAACACGT	
Db	781 AGAATTGTGTTGGAGGAAATGTATTAACACGT	
Qy	841 ATTTCCTCGAAAATCAATAAGTTGCACTGTGCG	
Db	841 ATTTCCTCGAAAATCAATAAGTTGCACTGTGCG	
Qy	901 CCACGTAGAATGTGTTAAATGGCACTGTGCG	

Db	901	CCACGTAGATGATGTTAAATGGGACTTGCGTATTCTGCATATACTTAAATAA	960
Qy	961	TAAAAAGTCATCATGTTCAACATCCTTACATAAAAAGCC	1004
Db	961	TAAAAAGTCATCATGTTCAACATCCTTACATAAAAAGCC	1004
RESULT 2	AR167384	AR167384	1004 bp
LOCUS	Sequence 20	from patent US 6287569.	DNA
DEFINITION			linear
ACCESSION	AR167384		PAT 17-DEC-2001
VERSION	AR167384.1	GI:17903161	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1004)		
AUTHORS	Kipps, T.J. and Wu, Y.		
TITLE	Vaccines with enhanced intracellular processing		
JOURNAL	Patent: US 6287569-A 20 11-SEP-2001;		
FEATURES	Location/Qualifiers		
source	1..1004		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
	Query Match	Score 1004; DB 6; Length 1004;	
	Best Local Similarity	100.0%; Pred. No. 3.7e-223;	
	Matches 1004;	Mismatches 0; Indels 0; Gaps 0;	
Qy	1 CGCCATTATTGGGTTCCTGGTATCTCCCGTCTGGCTCTGGTCTCCGGCTTAGAGGGACC	60	
Db	1 CGCCATTATTGGGTTCCTGGTATCTCCCGTCTGGCTCTGGTCTCTGGCTTAGAGGGACC	60	
Qy	61 AGGAGAGGGGGAGTTGGAGCTGGAGGTAAACACCTGGGTCTCACTCTGGATG	120	
Db	61 AGGAGAGGGGGAGTTGGAGCTGGAGGTAAACACCTGGGTCTCACTCTGGATG	120	
Qy	121 GTGGTGGCAACAGAGATGGCAAGGGAGGTGGAGTTAGGGGGCTGAGGGTAGG	180	
Db	121 GTGGTGGCAACAGAGATGGCAAGGGAGGTGGAGTTAGGGGGCTGAGGGTAGG	180	
Qy	181 AGTGGGGCTGGAGGAGTAAGATGGGGCAAGAGGGTCTGGTGTGAGGTGAGGCTG	240	
Db	181 AGTGGGGCTGGAGGAGTAAGATGGGGCAAGAGGGTCTGGTGTGAGGTGAGGCTG	240	
Qy	241 TGCTCCAAGCAGGTGATGAAAGGAGTCCCCTGGTGTGAGGTGAGGCTG	300	
Db	241 TGCTCCAAGCAGGTGATGAAAGGAGTCCCCTGGTGTGAGGTGAGGCTG	300	
Qy	301 AAGACGGCAAGCTGTGGTTCACTTCTGAGTTGTGGCAGCACCTGTGGAGCG	360	
Db	301 AAGACGGCAAGCTGTGGTTCACTTCTGAGTTGTGGCAGCACCTGTGGAGCG	360	
Qy	361 GCAGGTCAAAGGAGCAATTAGGGAGGTTCACTGTGAGGTGGAGGATGCTTC	420	
Db	361 GCAGGTCAAAGGAGCAATTAGGGAGGTTCACTGTGAGGTGGAGGATGCTTC	420	
Qy	421 GATCCTGACTCTGATCGCCGCCCTGGCTTCCAAAGTGGCGAGATTACAGCGAT	480	
Db	421 GATCCTGACTCTGATCGCCGCCCTGGCTTCCAAAGTGGCGAGATTACAGCGAT	480	
Qy	481 GTGCAATTGTAAAGGACTTTGGGACCACTATCAAAATGCTGTGAAGAGAAATGCTACCCAGA	540	
Db	481 GTGCAATTGTAAAGGACTTTGGGACCACTATCAAAATGCTGTGAAGAGAAATGCTACCCAGA	540	
Qy	541 TGTATCATTATCCTGCTGCTGAGGAGCGCTCTTCTAGGATTCTGTCACTCTTC	600	
Db	541 TGTATCATTATCCTGCTGCTGAGGAGCGCTCTTCTAGGATTCTGTCACTCTTC	600	
Qy	601 TGCTTGTCCAGAACACATGACCAAGCTGCAATTACGCTTAACCTTAACTACGCAATAC	660	

Db	601	TGCTTTGGTCAGAACATATTGACCAAGGCTCCTGAAAGTGTAGTTACTAAGCATAGAC	660	Qy	301	AAGACCGCAAGCTGTGTGTTCACTTCTGAGTTGTCAGGTCAGCGTGTGAGACGAGC	360
Qy	661	TTTTAAACTTCAACCAGTATTACTGAAATAAACAAATGTTGAAATTCCCTGAGTGT	720	Db	301	AAGACGGCAAGGAGCAATTAGGGAGGATGAGGATGTTCACTTCTGAGTTGTCAGGAGC	360
Db	661	TTTAAACTTCAACCAGTATTACTGAAATAAACAAATGTTGAAATTCCCTGAGTGT	720	Qy	361	GCAGGCTCAAGGAGCAATTAGGGAGGATGAGGATGTTCACTTCTGAGTTGTCAGGAGC	420
Qy	721	TATCTCTGTATTAAAGGTAATAAACATATCATTAATCTGGGATCATGCC	780	Db	361	GCAGGCTCAAGGAGCAATTAGGGAGGATGAGGATGTTCACTTCTGAGTTGTCAGGAGC	420
Db	721	TATCTCTGTATTAAAGGTAATAAACATATCATTAATCTGGGATCATGCC	780	Qy	421	GATCTCTGACCTCTGATCGCCGCCCTGGCTTCCAAAGTCGAGATTACAGCGAT	480
Qy	781	AGAGATTGTTGGGAGGGAAATGTTACAGGTTTCTGAAATTAAATCCAAAGTT	840	Db	421	GATCTCTGACCTCTGATCGCCGCCCTGGCTTCCAAAGTCGAGATTACAGCGAT	480
Db	781	AGAGATTGTTGGGAGGGAAATGTTACAGGTTTCTGAAATTAAATCCAAAGTT	840	Qy	481	GTGCAATTGTAAGGACTCTGGGACCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA	540
Qy	841	ATTTCCTCAGAAAATCAAAATAAGTTGTCATGTTTATCTTAAACATTAAAAA	900	Db	481	GTGCAATTGTAAGGACTCTGGGACCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA	540
Db	841	ATTTCCTCAGAAAATCAAAATAAGTTGTCATGTTTATCTTAAACATTAAAAA	900	Qy	541	TGTATCAATTACCTCTGCTGAGGCGCTCTTTAGGATTCACTCACACTCTGC	600
Qy	901	CCACTGTGAAATGATGTTAAATAGGGACTGTGGCAGTATTCTGACATACTATAAAATA	960	Db	541	TGTATCAATTACCTCTGCTGAGGCGCTCTTTAGGATTCACTCACACTCTGC	600
Db	901	CCACTGTGAAATGATGTTAAATAGGGACTGTGGCAGTATTCTGACATACTATAAAATA	960	Qy	601	TGCTTGTCAGAAACAATGACCAAGCTCTGAAAGATGTAAGTTACTAGATAAGC	660
Qy	961	TTAAAAAGTCATAGTATTCAACATCTTTAACACTAAAAGCC	1004	Db	601	TGCTTGTCAGAAACAATGACCAAGCTCTGAAAGATGTAAGTTACTAGATAAGC	660
Db	961	TTAAAAAGTCATAGTATTCAACATCTTTAACACTAAAAGCC	1004	Qy	661	TTTAAACTTCAACATGATTGATTTGAAATAATGTAATTCCTGAGTGT	720
Db	661	TTTAAACTTCAACATGATTGATTTGAAATAATGTAATTCCTGAGTGT	720	Db	721	TATTGCACTGTATTAAAAGCTAAATAACATCAAAATGAGGATCTTTGAGC	780
Qy	721	TATTGCACTGTATTAAAAGCTAAATAACATCAAAATGAGGATCTTTGAGC	780	Db	721	TATTGCACTGTATTAAAAGCTAAATAACATCAAAATGAGGATCTTTGAGC	780
DEFINITION	Sequence 1 from patent US 6465184.	1004 bp	DNA	LINEAR	PAT 20-DEC-2002		
ACCESSION	AR236412						
VERSION	AR236412						
KEYWORDS	GI:27280387						
SOURCE	Unknown.						
ORGANISM	Unclassified. 1 (bases 1 to 1004)						
REFERENCE	Boel,P., Wildmann,C., Boon-Falleur,T., van der Bruggen,P., Coulie,P. and Renaud,J.-C.						
AUTHORS							
DEFINITION	Isolated nucleic acid molecule encoding peptides which form complexes with MHC molecules HLA-Cw*1601 and uses thereof						
TITLE							
JOURNAL	Patent: US 6465184-A 11-OCT-2002;						
FEATURES	Location/Qualifiers						
source	1.. 1004 /organism="unknown" /mol_type="genomic DNA"						
ORIGIN							
Query Match	100.0%	Score 1004; DB 6; Length 1004;					
Best Local Similarity	100.0%	Prod. No. 3..7e-223;					
Matches 1004; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;			
Qy	1	CGCCATTAGGGCTTCGGTATCTCCCGCTGAGCTCTGTTCCCGCTTAGAGAAC	60	AR411457	AR411457	1004 bp	DNA
Db	1	CGCCATTAGGGCTTCGGTATCTCCCGCTGAGCTCTGTTCCCGCTTAGAGAAC	60	LOCUS DEFINITION	Sequence 1 from patent US 6638512.		
Qy	61	AGGAGAAGGGGGAGGTGGAGCTGGAGCTGGAGCTGTAACACCGTGTCACTCTGAGT	120	VERSION	AR411457	linear	PAT 18-DEC-2003
Db	61	AGGAGAAGGGGGAGGTGGAGCTGGAGCTGGAGCTGTAACACCGTGTCACTCTGAGT	120	KEYWORDS	GI:40163669		
Qy	121	GTGGTGGCAACAGAGATGGCAGGGCAGCTGGAGCTGGAGGCCCCCTGAGGGTAGG	180	ORGANISM	Unknown.		
Db	121	GTGGTGGCAACAGAGATGGCAGGGCAGCTGGAGCTGGAGGCCCCCTGAGGGTAGG	180	REFERENCE	1 (bases 1 to 1004)		
Qy	181	AGTGGGGCTGGAGGCTTAACATGGCGCAAGGGTTTCTGGCATGTCAGGCTGC	240	AUTHORS	Boel,P., Wildmann,C., Boon-Falleur,T., van der Bruggen,P., Coulie,P. and Renaud,J.-C.		
Db	181	AGTGGGGCTGGAGGCTTAACATGGCGCAAGGGTTTCTGGCATGTCAGGCTGC	240	TITLE	Method for treating subjects by using BAGE tumor rejection antigens precursors or tumor rejection antigens		
Qy	241	TGCTTCAAGCCAGGTGTAAGTGGAGGATCAAGGGTTTCTGGCATGTCAGGCTG	300	JOURNAL	Patent: US 6638512-A 1 28-OCT-2003;		
Db	241	TGCTTCAAGCCAGGTGTAAGTGGAGGATCAAGGGTTTCTGGCATGTCAGGCTG	300	FEATURES	Location/Qualifiers		
Qy	241	TGCTTCAAGCCAGGTGTAAGTGGAGGATCAAGGGTTTCTGGCATGTCAGGCTG	300	/organism="unknown"			
Db	241	TGCTTCAAGCCAGGTGTAAGTGGAGGATCAAGGGTTTCTGGCATGTCAGGCTG	300	/mol_type="genomic DNA"			
Query Match	100.0%	Score 1004; DB 6; Length 1004;					
Best Local Similarity	100.0%	Pred. No. 3..7e-223;					

Match ID	Text	Score	Start	End	Length	Method	Organism	Key
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;								
1	CGCCATTAGGTCTCCGATATCPCCGTGTAGCTGGCTTGAGCTGGAC	60	0	60	60	BD176601	Homo sapiens (human)	Method of molecular diagnosis of chronic myelogenous leukemia.
1	CGCCATTAGGTCTCCGATATCPCCGTGTAGCTGGCTTGAGCTGGAC	60	0	60	60	BD176601	Bivalvia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Hominidae; Homo.	Method of molecular diagnosis of chronic myelogenous leukemia.
61	AGGAGAAGGGAGCTGAACTGAGCTGTAACACCTGGCTCACTCTGGAT	120	0	120	120	BD176601	Homo sapiens (human)	Method of molecular diagnosis of chronic myelogenous leukemia.
61	AGGAGAAGGGAGCTGAGCTGAACTGAGCTGTAACACCTGGCTCACTCTGGAT	120	0	120	120	BD176601	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Hominidae; Homo.	Method of molecular diagnosis of chronic myelogenous leukemia.
121	GTCGTCGCCAACAGAGATGGCGAGCTGGCTGAGGGGGCTGAGCTGGAG	180	0	180	180	BD176601	Homo sapiens (human)	Method of molecular diagnosis of chronic myelogenous leukemia.
121	GTCGTCGCCAACAGAGATGGCGAGCTGGCTGAGGGGGCTGAGCTGGAG	180	0	180	180	BD176601	Mamo. H.	Method of molecular diagnosis of chronic myelogenous leukemia.
181	AGTGGGACTGAGCAATAGGGGGCAAGAGCGTTTTCGGCATGTCGCCAGC	240	0	240	240	BD176601	OS Homo sapiens (human)	Method of molecular diagnosis of chronic myelogenous leukemia.
181	AGTGGGACTGAGCAATAGGGGGCAAGAGCGTTTTCGGCATGTCGCCAGC	240	0	240	240	BD176601	PN	Method of molecular diagnosis of chronic myelogenous leukemia.
241	TGCTCCAAGCCAGGCTGAACTGAGCTGGCTGAGGGTTGGAGCTG	300	0	300	300	BD176601	PR	Method of molecular diagnosis of chronic myelogenous leukemia.
241	TGCTCCAAGCCAGGCTGAACTGAGCTGGCTGAGGGTTGGAGCTG	300	0	300	300	BD176601	PI	Method of molecular diagnosis of chronic myelogenous leukemia.
301	AAGACGCCAACGCTGTCGTTCACTTCTGAGGTGTCGGCTGAGCTGAGC	360	0	360	360	BD176601	PC	Method of molecular diagnosis of chronic myelogenous leukemia.
301	AAGACGCCAACGCTGTCGTTCACTTCTGAGGTGTCGGCTGAGCTGAGC	360	0	360	360	BD176601	FT	Method of molecular diagnosis of chronic myelogenous leukemia.
361	GCAGCTCAACAGGCCAATAGGAGGAGATGGAGTTCACTGTGTAGCGAGGTC	420	0	420	420	BD176601	FT	Method of molecular diagnosis of chronic myelogenous leukemia.
361	GCAGCTCAACAGGCCAATAGGAGGAGATGGAGTTCACTGTGTAGCGAGGTC	420	0	420	420	BD176601	FEATURES	Location/Qualifiers
421	GATCTCTGACTCTCGCTGATCGCCGCGCTTCCAAAGTCGCCGAGATTACAGGGAT	480	0	480	480	BD176601	source	1..1004
421	GATCTCTGACTCTCGCTGATCGCCGCGCTTCCAAAGTCGCCGAGATTACAGGGAT	480	0	480	480	BD176601	source	1..1004
481	GTGCATTTGTAAAGCACTTGGGCCACATTCATAATGCTGTGAGAGAAATGCCAGA	540	0	540	540	BD176601	ORIGIN	1..1004
481	GTGCATTTGTAAAGCACTTGGGCCACATTCATAATGCTGTGAGAGAAATGCCAGA	540	0	540	540	BD176601	Query Match	100.0%
541	TGATCATATATCCTCTGTCAGGAGCTCTGAGGTTCACTCATCTTC	600	0	600	600	BD176601	Best Local Similarity	100.0%
541	TGATCATATATCCTCTGTCAGGAGCTCTGAGGTTCACTCATCTTC	600	0	600	600	BD176601	Matches	1004;
601	TGCTTGTCGAAACATGAGCAAGCTGAGGTTCACTCATCTTC	660	0	660	660	BD176601	Pred. No.	3..7e-223;
601	TGCTTGTCGAAACATGAGCAAGCTGAGGTTCACTCATCTTC	660	0	660	660	BD176601	Mismatches	0;
661	TTTTAAACTCAACCAATGTTATTACTGAAAATACAATGTTGTAATTCCCTGAGTT	720	0	720	720	BD176601	Indels	0;
661	TTTTAAACTCAACCAATGTTATTACTGAAAATACAATGTTGTAATTCCCTGAGTT	720	0	720	720	BD176601	Gaps	0;
721	TATTCTACTGTATTAAGGTTAAATCATATACTAAATCTGGGGATCATGCC	780	0	780	780	BD176601	KEY	Method of molecular diagnosis of chronic myelogenous leukemia.
721	TATTCTACTGTATTAAGGTTAAATCATATACTAAATCTGGGGATCATGCC	780	0	780	780	BD176601	ORGANISM	Homo sapiens (human)
781	AGAGATGTTGGGGAGGAGATGTTCACTGGTTCACTGAACTATTAATAA	840	0	840	840	BD176601	REFERENCE	Method of molecular diagnosis of chronic myelogenous leukemia.
781	AGAGATGTTGGGGAGGAGATGTTCACTGGTTCACTGAACTATTAATAA	840	0	840	840	BD176601	AUTHORS	Mamo. H.
841	ATTCTCTGAGAAATCAATAAAGTGGATGTTTATCTTAACTATTAATAA	900	0	900	900	BD176601	TITLE	Method of molecular diagnosis of chronic myelogenous leukemia.
841	ATTCTCTGAGAAATCAATAAAGTGGATGTTTATCTTAACTATTAATAA	900	0	900	900	BD176601	JOURNAL	FUJISAWA PHARMACEUTICAL CO LTD, HIROYUKI MAMO
901	CCACTGTAAGTGTAAATAGGGACTCTGGCAATTCTGACATATAAAATA	960	0	960	960	BD176601	COMMENT	Method of molecular diagnosis of chronic myelogenous leukemia.
901	CCACTGTAAGTGTAAATAGGGACTCTGGCAATTCTGACATATAAAATA	960	0	960	960	BD176601	CDS	Method of molecular diagnosis of chronic myelogenous leukemia.
961	TTRAAAGTCATCTGATTCACATCTTACACTAAAGGCC	1004	0	1004	1004	BD176601	PR	Method of molecular diagnosis of chronic myelogenous leukemia.
961	TTRAAAGTCATCTGATTCACATCTTACACTAAAGGCC	1004	0	1004	1004	BD176601	PI	Method of molecular diagnosis of chronic myelogenous leukemia.
541	TGTATGTTATTCCTGTCGTTGCTGAGGAGCCGGCTCTTCAGGATTTCAGTCATCTTC	600	0	600	600	BD176601	PC	Method of molecular diagnosis of chronic myelogenous leukemia.









**Sario, A.D.**  
New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile  
*Eur. J. Hum. Genet.* 10 (12), 833-840 (2002)

**JOURNAL** EUR. J. HUM. GENET.  
**MEDLINE ID** 22349465  
**PUBMED ID** 12461691  
**REFERENCE** 2 (bases 1 to 1068)  
Rauault, M., Ventura, M., Galtier, N., Brun, M.-E., Archidiacono, N., Roizes, G. and De Sario, A.

**TITLE** BAGE, a gene family generated by juxtacentromeric reshuffling in the Hominidae lineage, is under selective pressure  
**JOURNAL** *Genomics* (2003) In press

**REFERENCE** 3 (bases 1 to 1068)  
Rauault, M., van der Bruggen, P., Brun, M.-E., Boyle, S., Roizes, G. and De Sario, A.

**TITLE** Direct Submission  
**JOURNAL** Submitted (08-JUL-2002) Institut de Gentique Humaine, CNRS UPR 1142, 141, rue de la Cardinale, Montpellier 34395, France  
**FEATURES** Location/Qualifiers  
1. .1068  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:3606"  
/chromosome="13"  
1.  
/gene="BAGE1"  
/gene\_id="201..332"  
/note="umor antigen; alternatively spliced"  
/codon\_start=1  
/product="B melanoma antigen variant e"  
/protein\_id="AAQ12673..1"  
/db\_xref="GI:28173036"  
/translation="MAAKAVFLNLSAQILQLRMLKESPSVWRLSPEDGTALCFIF"

**ORIGIN**

Query Match 56.1%; Score 563.4; DB 9; Length 1068;  
Best Local Similarity 90.5%; Pred. No. 1.5e-120;  
Matches 647; Conservative 0; Mismatches 11; Indels 57; Gaps 2;

Db 1 CGCCATTAGGTTCTCGTATTCGCCCTGACTCTGCTGTTCCCGCTTAGAGACC 60  
1 CGCCATTAGGTTCTCGTATTCGCCCTGACTCTGCTGTTCCCGCTTAGAGACC 60

QY 61 AGGAGAGGGGGAGCTGGAGCTGGACCTGGCTGTAACCCCTGGCTCACTCTGGATG 120  
61 AGGAGAGGGGGAGCTGGAGCTGGACCTGGCTGTAACCCCTGGCTCACTCTGGATG 120

QY 121 GTGCTGCAACAGAGATGGCAGCTGGAGCTGGAGCTGGCTGAGCGGTAGG 180  
121 GTGCTGCAACAGAGATGGCAGCTGGAGCTGGAGCTGGCTGAGCGGTAGG 180

Db 181 AGTGGGGCTGAGAGTAAGATGGCCCAAGGGCTTTCGGCATGTCGCCAAG 240  
181 AGTGGGGCTGAGAGTAAGATGGCCCAAGGGCTTTCGGCATGTCGCCAAG 240

Db 241 TGCTCAAGCAGGGTGTGATCTGCCTGAGCTGGCTGAGTGGGGTGGCTG 300  
241 TGCTCAAGCAGGGTGTGATCTGCCTGAGCTGGCTGAGTGGGGTGGCTG 300

QY 301 AAGACCGCAAGCTCTGCTGAGCTGGCACCCTGATGGCTGAGCTGGGGTGGCTG 360  
301 AAGACCGCAAGCTCTGCTGAGCTGGCACCCTGATGGCTGAGCTGGGGTGGCTG 360

QY 361 GCAGCTCAAAGGGCAAATAGGAG-----GAGTGGAGTTCACT 400  
361 GCAGCTCAAAGGGCAAATAGGAGTAACGGCAAGTGGAGTTCACT 420

Db 401 GTGTCAGGCCAGGATGTTCTGATCTGCCTGAGCTGGCTGAGCTGGCTG 460  
421 GTGTCAGGCCAGGATGTTCTGATCTGCCTGAGCTGGCTGAGCTGGCTG 480

QY 461 AGTGGGGAGATA-----CAGGGATGTC 483  
Db 481 ACTGCCGAGATTACAGCTTATGAAAGAGATTCTCCCTTACGGATGTC 540

QY 484 CATTGTGAACTTGGACCTCATCAATGCTGAGAGAAATGACCCAGATG 543  
Db 541 CATTGTGAACTTGGACCTCATCAATGCTGAGAGAAATGACCCAGATG 603

QY 544 ATCATPATCCCTGTCGAGGGGGCTCTTGAGATTCTGAGTCACTCTCTGC 603  
Db 601 ATCATPATCCCTGTCGAGGGGGCTCTTGAGATTCTGAGTCACTCTCTGC 660

QY 604 TTGTGTCAGAACACATTGACCAAGCTCTGAAGGTGAAAGTTACTACCGATAG 658  
Db 661 TTGTGTCAGAACACATTGACCAAGCTCTGAAGGTGAAAGTTACTACCGATAG 715

**RESULT** 12  
**ACCESSION** AC064811/c  
**LOCUS** AC064811  
**DEFINITION** Homo sapiens chromosome 4 clone RP1-63BN24 map 4, WORKING DRAFT  
**SEQUENCE** AC064811  
Homo sapiens chromosome 4, clone RP1-63BN24  
**VERSION** AC064811.2  
**HTGS\_PHRSE1**; HTGS\_DRAFT.  
**KEYWORDS**  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**EBI** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrini; Homiridae; Homo.  
**REFERENCES** 1 (bases 1 to 177896)  
**AUTHORS** Birren, B.; Linton, L.; Nusbaum, C. and Lander, E.  
**TITLE** Homo sapiens chromosome 4, clone RP1-63BN24  
**JOURNAL** Unpublished  
**AUTHORS** Birren, B.; Linton, L.; Nusbaum, C., Lander, E., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavskiy, L., Boukhalter, B., Brown, A., Burlett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Coote, P., Dereillano, K., Devar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferrera, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Govertse, M., Graham, L., Grand-Pierre, N., Grant, G., Hago, B., Heatord, J., Jones, C., Kann, L., Karatas, A., Howland, J.C., Iliev, I., Johnson, R., Klein, J., Larocque, K., Lamazares, R., Landers, T., Lebovky, J.N., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Melidrim, J., Menes, L., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterkin, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spender, B., Stange-Thomann, N., Stoianovic, N., Subramanian, A., Talamas, J.J., Testay, S., Theodore, J., Terrell, A., Travers, M., Trigilio, J., Vassiliev, H., Vie, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (22-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**COMMENT** On May 22, 2000 this sequence version replaced gi:7637317. All repeats were identified using RepeatMasker.  
Smit, A.P.A. & Green, P. (1996-1997)  
**HTTP** http://ftp.genome.washington.edu/RM/RepeatMasker.html  
**GENOME CENTER** Genome Center  
**CENTER** Whitehead Institute / MIT Center for Genome Research  
**CENTER CODE** WIBR  
**WEB SITE** http://www-seq.wi.mit.edu  
**CONTACT** sequence\_submissions@genome.wi.mit.edu  
**PROJECT INFORMATION**  
**CENTER PROJECT NAME** 169335  
**CENTER CLONE NAME** 638\_N\_24  
**SUMMARY STATISTICS**  
**SEQUENCING VECTOR** M13 M77815; 100% of reads  
**CHEMISTRY** Dye-terminator Big Dye; 100% of reads





RESULT 14  
AC145613

LOCUS AC145613 chromosome 17830 bp DNA LINEAR FTG 19-JUL-2003  
DEFINITION Homo sapiens chromosome UNK clone RP1-1266H24, \*\*\* SEQUENCING IN  
PROGRESS \*\*, 14 unordered pieces.

ACCESSION AC145613

VERSION GI:32996952

KEYWORDS MHC; HRC; PHASE1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1. (bases 1 to 178300) Wilson, R.K.

AUTHORS Wilson, R.K.

TITLE The sequence of Homo sapiens clone  
Unpublished  
2. (bases 1 to 178300)

REFERENCE Wilson, R.K.

AUTHORS Direct Submission

TITLE Submitted (19-JUL-2003) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA

COMMENT -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://Genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H\_NH1266H24  
----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Chemistry: Dye-primer ET; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 171980 bases at least Q40  
Consensus quality: 173066 bases at least Q20

FEATURES

source /organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="UNK"  
/clone="RP11-1266H24"

1..1534 misc\_feature /note="assembly\_name:Contig13"  
1635..4097 misc\_feature /note="assembly\_name:Contig14"  
4198..6161 misc\_feature /note="assembly\_name:Contig15"  
6262..11394 misc\_feature /note="assembly\_name:Contig16"  
11495..11331 misc\_feature /note="assembly\_name:Contig17"  
13412..16041 misc\_feature /note="assembly\_name:Contig18"  
16142..21041 misc\_feature /note="assembly\_name:Contig19"  
26589..30637 misc\_feature /note="assembly\_name:Contig20"  
30793..34020 misc\_feature /note="assembly\_name:Contig21"  
31850..38017 misc\_feature /note="assembly\_name:Contig22"  
31850..34020 misc\_feature /note="assembly\_name:Contig23"  
31850..38017 misc\_feature /note="assembly\_name:Contig24"  
31850..38017 misc\_feature /note="assembly\_name:Contig25"  
31850..38017 misc\_feature /note="assembly\_name:Contig26"  
31850..38017 misc\_feature /note="assembly\_name:Contig27"  
31850..38017 misc\_feature /note="assembly\_name:Contig28"

ORIGIN

Query Match 52.4%; Score 526; DB 2; Length 178300;  
Best Local Similarity 99.1%; Pred. No. 9.8-12;  
Matches 529; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 471 TTACCGATGTGCATTGTAACTTGGCCACTCATCAATGCTGTGAAGAGAA 530  
Db 165850 TTTACCGATGTGCATTGTAACTTGGCCACTCATCAATGCTGTGAAGAGAA 165909

Qy 531 TGTACCCAGATGATGATAATTATCCCTGTGAGAGCCGGCTCTTCAGATTCTG 590  
Db 165910 TGTACCCAGATGATGATAATTATCCCTGTGAGAGCCGGCTCTTCAGATTCTG 165969

Qy 591 CACATCTTCCTGCTTGTCCAGAACATGACCAAAGCTCTGAAAGATGTAAGTTACT 650

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 14 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1534: contig of 1534 bp in length  
\* 1535 1634: gap of unknown length  
\* 1635 4097: contig of 2463 bp in length  
\* 4098 4197: gap of unknown length

Db	165070	CACATCTCCGCTTGTCCCGAAAACATTGACCAAGCTCTGAAGTGTACTTACT	166029	Qy	531	TGTACCCAGATGTATTATCCTTGTGCGAGGAGCCGGCTCCTTCAGGATTTCACT	590
Qy	651	ACGCATAGACTTTAACCTCAACCAATGATTACTGAAATAACAAATGTTAAATT	710	Db	77638	TGTACCCAGATGTATTATCCTTGTGCGAGCTCTGAAGTGTACT	77629
Db	166030	ACGCATAGACTTTAACCTCAACCAATGATTACTGAAATAACAAATGTTAAATT	166089	Qy	591	CACATCTCCGCTTGTCCAGAACACATTGACCAAGCTCTGAAGTGTACTTACT	650
Qy	711	CCCTGAGGTATTCTACTGTATTAAAGTAAATAATCATTTAACTGAGGG	770	Db	77628	CACATCTCCGCTTGTCCAGAACACATTGACCAAGCTCTGAAGTGTACTTACT	77569
Db	166090	CCCTGAGGTATTCTACTGTATTAAAGTAAATAATCATTTAACTGAGGG	166149	Qy	651	ACCATAGACTTTAACCTCAACCATGATTAACTGAAATAAAATGTTAAATT	710
Qy	771	GATCATTCGGAGGATTGTGGGAGGAATGTTCAACGGTTCTATTGAAATTAAAT	830	Db	77568	ACCGATAGACTTTAACCTCAACCATGATTAACTGAAATAAAATGTTAAATT	77509
Db	166150	GATCGTTGCCAGAGTTGTGGGGGGAGAATGTTCAACGGTTCTATTGAAATTAAAT	166209	Qy	711	CCCTGAGGTATTCTACTGTATTAAAGTAAATACTTAACTGAGG	770
Qy	831	CCAAAAGGTATTCTCCAGAAAATCAAAATAACTTGCATGTTTATTCTTAAAC	890	Db	77508	CCCTGAGGTATTCTACTGTATTAAAGTAAATACTTAACTGAGG	77449
Db	166210	CCAAAAGGTATTCTCCAGAAAATCAAAATAAGTTCATGTTTATTCTTAAAC	166269	Qy	771	GATCATTCGGAGGATTGTGGGAGGAATGTTCAACGGTTCTATTGAAATTAAAT	830
Qy	891	ATTTAAAAACCACGTGAGATGATGAAATAGGGACTGTGCAGTTCTGCATATAC	950	Db	77448	GATCGTTGCCAGAGTTGTGGGGAGAATGTTCAACGGTTCTATTGAAATTAAAT	77389
Db	166270	ATTTAAAAACCACGTGAGATGAAATAGGGACTGTGCAGTTCTGCATATAC	166329	Qy	831	CCAAAAGGTATTCTCCAGAAAATCAAAATAAGTTCATGTTTATTCTTAAAC	890
Qy	951	TATAAAATTATAAAAGTCAATCGTATTCACATCTTAACTAAAGGCC	1044	Db	77388	CCAAAAGGTATTCTCCAGAAAATCAAAATAAGTTCATGTTTATTCTTAAAC	77329
Db	166330	TATAAAATTATAAAAGTCAATCGTATTCACATCTTAACTAAAGGCC	166383	Qy	891	ATTATAAAACCACTTAGATGTGTTAACTAGGACTGTGCAGATTCTGACATATAC	950
Db	77328	ATTATAAAACCACTTAGATGTGTTAACTAGGACTGTGCAAATTCTGACATATAC	77269	Db	77328	ATTATAAAACCACTTAGATGTGTTAACTAGGACTGTGCAAATTCTGACATATAC	77269
RESULT 15				Qy	951	TATAAAATTATAAAAGTCAATCGTATTCACATCTTAACTAAAGGCC	1004
AF155875/C	AF155875	107701 bp DNA linear HTG 31-DEC-2001		Db	77268	TATAAAATTATAAAAGTCAATCGTATTCACATCTTAACTAAAGGCC	77215
LOCUS				Search completed: July 9, 2004, 08:57:30			
DEFINITION	Homo sapiens clone MC169,	*** SEQUENCING IN PROGRESS ***, 2 ordered pieces		Job time : 6481 secs			
ACCESSION	AF155875						
VERSION	AF155875.1	GI:18000259					
KEYWORDS	HTGS; PHASE2.						
SOURCE	Homo sapiens (human)						
ORGANISM							
REFERENCE							
AUTHORS							
TITLE	Xia,J.-H., Huang,L., Tang,D.-S., Dai,H.-P., Pan,Q. and Long,Z.-G.						
JOURNAL	Direct Submission						
	Submitted (18-MAY-1999) National Lab of Medical Genetics of China, Hunan Medical University, 88 Xiang ya Road, Changsha, Hunan 410078, P.R. China						
COMMENT	* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.						
	* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.						
	* are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.						
	* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.						
	* are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.						
	* are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.						
FEATURES							
source	1..107701	/organism="Homo sapiens"					
		/mol_type="Genomic DNA"					
		/db_xref="taxon:9606"					
		/clone="MC169"					
ORIGIN							
Query Match	52.2%	Score 524.4; DB 2; Length 107701;					
	Best Local Similarity 98.9%; Pred. No. 2,3e-111;						
	Matches 526; Conservative 0; Mismatches 6; Indels 0; Gaps 0;						
Qy	471	TTACAGCGATGTGCATTGTAAAGCAGTTGGACCTCAATGCTGTGAAGACAA	530				
Db	77748	TTTCAGCAATGTGCATTGTAAAGCAGTTGGACCTCAATGCTGTGAAGACAA	77689				

This Page Blank (uspto)



WPI; 1996 393411/39.	721	TATTCCTACTGTAAAGGTAATAACATACTTAAATCTGAGGGATCATGGCC	780
P-PPDB; AAW02152.	721	TATTCCTACTGTAAAGGTAATAACATACTTAAATCTGAGGGATCATGGCC	780
Tumour rejection antigen precursor (TRAP) and gene - useful to develop pros. for diagnosis and treatment of disorders characterised by TRAP, partic. melanomas.	781	AGAGATGTGGGGAGGAATGTTACCGGTTCATGAAATTAAATCCAAAAGTT	840
XX	781	AGAGATGTGGGGAGGAATGTTACCGGTTCATGAAATTAAATCCAAAAGTT	840
Claim 1; Page 27-28; 44pp; English.	Db		
XX	Db		
A cDNA clone (AAT16382), designated cDNA-AD5, codes for the BAGE tumour rejection antigen precursor (TRAP) (AAW02152). MHC molecule HLA-Cw*1601 presents a tumour rejection antigen derived from the BAGE TRAP. The cDNA was identified by cotransfected HLA-Cw*1601 cDNA with a cDNA library derived from melanoma cell line MZ-MEL43 into COS-7 cells and isolating clones capable of stimulating proliferation of tumour necrosis factor by cytotoxic T-lymphocytes. In 600 samples of tumors, the BAGE gene was expressed mainly in melanomas (esp. in metastatic lesions), bladder carcinomas and primary carcinomas. The isolated gene may be used diagnostically or for prodn. of the BAGE TRAP	Qy	841 ATTTCCTCAGAAAATCAATAAGTTGATGTTTATTCTAAACATTTAAAA	900
PS	Qy	841 ATTTCCTCAGAAAATCAATAAGTTGATGTTTATTCTAAACATTTAAAA	900
PS	Db	841 ATTTCCTCAGAAAATCAATAAGTTGATGTTTATTCTAAACATTTAAAA	900
PS	Qy	901 CCACGTAGAATGATGTTAAATGGGACTGTGCATTTCTGACATATAAATTA	960
PS	Db	901 CCACGTAGAATGATGTTAAATGGGACTGTGCATTTCTGACATATAAATTA	960
PS	Qy	961 TTAAAGCTCATCATCACATTCAACATCTTACACTAAAGGCC 1004	1004
PS	Db	961 TTAAAGCTCATCATCACATTCAACATCTTACACTAAAGGCC 1004	1004
SQ			
Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;	RESULT 2		
Best Local Similarity 100.0% ; Score 1004; DB 2; Length 1004;	ID	ABQ76214 standard; DNA; 1004 BP.	
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	ID	ABQ76214	
Qy	1	CGCCATTAGGGTCCCGTATCTCCCTGACTCTGCTCTGCTCCCGCTTAGAGAAC	60
Db	1	CGCCATTAGGGTCCCGTATCTCCCTGACTCTGCTCCCGCTTAGAGAAC	60
Qy	61	AGGAGACGGAGCTGGGGCTGAGCTGTAAACCCGGCGTCACTCTGATG	120
Db	61	AGGAGACGGAGCTGGGGCTGAGCTGTAAACCCGGCGTCACTCTGATG	120
Qy	121	GTCGTCGCAACAGAGATGGAGCGCAGCTGGAGTTAGAGCGGGCTGACCGTAGG	180
Db	121	GTCGTCGCAACAGAGATGGAGCGCAGCTGGAGTTAGAGCGGGCTGACCGTAGG	180
Qy	181	AGTGGGGCTGAGCGTAGATGGGGCAAGGGCTTTCGGCATTTCTGGCCAGC	240
Db	181	AGTGGGGCTGAGCGTAGATGGGGCAAGGGCTTTCGGCATTTCTGGCCAGC	240
Qy	241	TGCTCCAAAGCAGCTCTGCTTCACTCTGAGGTGGCTGGAGCTG	300
Db	241	TGCTCCAAAGCAGCTCTGCTTCACTCTGAGGTGGCTGGAGCTG	300
Qy	301	AAGACCGCACAGCTCTGCTTCACTCTGAGGTGGCTGGAGACG	360
Db	301	AAGACCGCACAGCTCTGCTTCACTCTGAGGTGGCTGGAGACG	360
Qy	361	GCAGCTCAACAGGCAATAGGAGGAACTGGGACTTCACTGCTGGAGCT	420
Db	361	GCAGCTCAACAGGCAATAGGAGGAACTGGGACTTCACTGCTGGAGCT	420
Qy	421	GATCCCTGACCTCGTGAACGCCGCTTCCAAAGCTCCGAGTTACAGCT	480
Db	421	GATCCCTGACCTCGTGAACGCCGCTTCCAAAGCTCCGAGTTACAGCT	480
Qy	481	GTGCAATTGTAAGACTTGAGCACTPATCAATGTTGAGAGAATGACCCAGA	540
Db	481	GTGCAATTGTAAGACTTGAGCACTPATCAATGTTGAGAGAATGACCCAGA	540
Qy	541	TGTATCATATCCCTGCTGCTGAGCCGGCTCTTCTGAGTTCACTCATCTTC	600
Db	541	TGTATCATATCCCTGCTGCTGAGCCGGCTCTTCTGAGTTCACTCATCTTC	600
Qy	601	TGCTTAAACTCAACCAATGTTACTGAAATAACAAATGTTACTACCGATAGAC	660
Db	601	TGCTTAAACTCAACCAATGTTACTGAAATAACAAATGTTACTACCGATAGAC	660
Qy	661	TTTTAAACTCAACCAATGTTACTGAAATAACAAATGTTACTACCGATAGCT	720
Db	661	TTTTAAACTCAACCAATGTTACTGAAATAACAAATGTTACTACCGATAGCT	720

Qy	Db	Score	Length
Query Match		100.0%	DB 2;
Best Local Similarity		100.0%	Pred. No. 1. 8e-250;
Matches 1004;		0; Mismatches 0;	Indels 0; Gap
Qy	1	CGCCAATTAGGCTTCGGTATCTCCGGTAGCTGCTGTCTGTCCTGAGCTGCTGTTAGAGGA	
Db	1	CGCCAATTAGGCTTCGGTATCTCCGGTAGCTGCTGTCTGTCCTGAGCTGCTGTTAGAGGA	
Qy	61	AGGAAACGGAAACTGGAGGTGGAGCTGGAGGTGGAGCTGGAGGTGGAGCTGGAGGTGGAGCTGGAG	
Db	61	AGGAAACGGAAACTGGAGGTGGAGCTGGAGGTGGAGCTGGAGGTGGAGCTGGAGGTGGAGCTGGAG	
Qy	121	GTCGGCAACAGAGATGCCAGGGCTGGAGGTGGAGCTGGAGGTGGAGCTGGAGGTGGAGCTGGAG	
Db	121	GTCGGCAACAGAGATGCCAGGGCTGGAGGTGGAGCTGGAGGTGGAGCTGGAGGTGGAGCTGGAG	
Qy	181	AGTGGGCTGGGGCAAGTAAGATGGGGCAGAAGGGCTTTCCTGCATTTCTGCATTTCTGCATTTCTGCAT	
Db	181	AGTGGGCTGGGGCAAGTAAGATGGGGCAGAAGGGCTTTCCTGCATTTCTGCATTTCTGCATTTCTGCAT	
Qy	241	TGCTCCAGCCAGGCTTGATGAGGAGGAGTCCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	
Db	241	TGCTCCAGCCAGGCTTGATGAGGAGGAGTCCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	
Qy	301	AGAGCGAACAGETCTGGCTTATCTTCTGAGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAG	
Db	301	AGAGCGAACAGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAG	
Qy	361	GGAGCTTAACAGGAGCATAGAGGGAGATGGAGGTTCACTGTGAGCTGAGCTGAGCTGAGCTGAG	
Db	361	GGAGCTTAACAGGAGCATAGAGGGAGATGGAGGTTCACTGTGAGCTGAGCTGAGCTGAGCTGAG	
Qy	421	GATCTCCGACCTTCGTTGACCCGCCCCCTGACCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTT	
Db	421	GATCTCCGACCTTCGTTGACCCGCCCCCTGACCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTT	
Qy	481	GTCGATTTGTGACCACTTGTGAGCCCTATGAACTGTGAGGAAAGTGTACCGAAATGTACCGAA	
Db	481	GTCGATTTGTGACCACTTGTGAGCCCTATGAACTGTGAGGAAAGTGTACCGAAATGTACCGAA	
Qy	541	TGATCATTTATCTTGTGCTGAGGACCGGGTCTTCTGAGATTCTGACATCTTCTGAGATTCTGACATCTT	
Db	541	TGATCATTTATCTTGTGCTGAGGACCGGGTCTTCTGAGATTCTGACATCTTCTGAGATTCTGACATCTT	
Qy	601	TGCTTTCGAGACACATTGACAAAGCTCTGAAAGATGTAAGTTTCTACGCTAG	
Db	601	TGCTTTCGAGACACATTGACAAAGCTCTGAAAGATGTAAGTTTCTACGCTAG	
Qy	661	TTTAAACCTCACCAATGTTACTCTGAAATAACAAATGTTGTAATACTTCTCTGTT	
Db	661	TTTAAACCTCACCAATGTTACTCTGAAATAACAAATGTTGTAATACTTCTCTGTT	
Qy	721	TATTCRACTTGTATTAAAGGTTATAATACATACTTATAATCTGAGGGATCATTTG	
Db	721	TATTCRACTTGTATTAAAGGTTATAATACATACTTATAATCTGAGGGATCATTTG	
Qy	781	AGAGATGTTGGGGAGAAATGTTATACCGTTCACTGAAATTAACTCAAAG	
Db	781	AGAGATGTTGGGGAGAAATGTTATACCGTTCACTGAAATTAACTCAAAG	
Qy	841	ATTCCTCTGAGAAAATCAAAATAAGTTGATGTTTATCTCTTAAACATTTAA	
Db	841	ATTCCTCTGAGAAAATCAAAATAAGTTGATGTTTATCTCTTAAACATTTAA	
Qy	901	CCACTGTAAATGATGTAATAGGGACTGTGCGATTTCTGACATACTATAAAAT	
Db	901	CCACTGTAAATGATGTAATAGGGACTGTGCGATTTCTGACATACTATAAAAT	
Qy	961	TTAAAAAGTCATCACTTACACATCTTACACTAAAGGCC 1004	
Db	961	TTAAAAAGTCATCACTTACACATCTTACACTAAAGGCC 1004	

Qy	241	TGCTCCAGCCAGGGCTGATGAAAGGAGTCCCCTGTGAGCTGGTTGGAGCTG 300	XX	Unidentified.
Db	241	TGCTCCAGCCAGGGCTGATGAAAGGAGTCCCCTGTGAGCTGGTTGGAGCTG 300	XX	US2003118592-A1.
Qy	301	AACAGCAGCTGCACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360	XX	26-JUN-2003.
Db	301	AACAGCAGCTGCACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360	XX	25-JUL-2002; 2002US-0C207655.
Qy	361	GCAGCTAACAGGAGCATAGGAGATGGAGATGGAGTTCACTCTGTCAGCAGGATGTC 420	XX	17-JAN-2001; 2001US-0367358P.
Db	361	GCAGCTAACAGGAGCATAGGAGATGGAGATGGAGTTCACTCTGTCAGCAGGATGTC 420	XX	17-JAN-2002; 2002US-000533530.
Qy	421	GATGTCGACCCCGTGATGCCCTGGCCCTAACACTGCCAATTACACGAT 480	XX	03-JUN-2002; 2002US-0385691P.
Db	421	GATGTCGACCCCGTGATGCCCTGGCCCTAACACTGCCAATTACACGAT 480	XX	(GENE-) GENECAST INC.
Qy	421	GATGTCGACCCCGTGATGCCCTGGCCCTAACACTGCCAATTACACGAT 480	XX	Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
Db	421	GATGTCGACCCCGTGATGCCCTGGCCCTAACACTGCCAATTACACGAT 480	XX	PA
Qy	481	GTCGATTGTAAGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540	XX	Hayden-Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
Db	481	GTCGATTGTAAGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540	XX	WPI; 2003-801317/75.
Qy	541	TGATCATATACTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600	XX	New binding domain-immunoglobulin fusion protein, useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, Grave's disease or autoimmune disease.
Db	541	TGATCATATACTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600	XX	XX
Qy	601	TGCTTGTCAGACACATTGCAAGCTTGTAAAGTAGTTACTACGGCATGAC 660	XX	PS Disclosure: SEQ ID NO 80; 157pp; English.
Db	601	TGCTTGTCAGACACATTGCAAGCTTGTAAAGTAGTTACTACGGCATGAC 660	XX	DR
Qy	661	TTTAAATTCATCAACAACTGAAATGTTAAATCCCTGAGT 720	XX	Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;
Db	661	TTTAAATTCATCAACAACTGAAATGTTAAATCCCTGAGT 720	XX	PT
Qy	721	TATCTCTGTATAAAGGTTAAATCATATAAAATCTGGGGATCATGCC 780	XX	Query Match 100.0%; Score 1004; DB 9; Length 1004;
Db	721	TATCTCTGTATAAAGGTTAAATCATATAAAATCTGGGGATCATGCC 780	XX	Best Local Similarity 100.0%; Prcd. No. 1; seq=250;
Qy	781	AGAGATTGTTGGGAGGAATGTTATCAACGGTTCTTGTAAATCAAAGTT 840	XX	Indels 0; PT
Db	781	AGAGATTGTTGGGAGGAATGTTATCAACGGTTCTTGTAAATCAAAGTT 840	XX	0; Gaps 0;
Qy	841	ATTCTCTGAGAAATCAAATAGTTGCAATTTTATCTAAACATTAAATAA 900	XX	CC Unidentified
Db	841	ATTCTCTGAGAAATCAAATAGTTGCAATTTTATCTAAACATTAAATAA 900	XX	Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;
Qy	901	CCACTGTAGATGATGAAATAGGACTGTGCAATTTCTGACATAATAATAA 960	XX	PT
Db	901	CCACTGTAGATGATGAAATAGGACTGTGCAATTTCTGACATAATAATAA 960	XX	Query Match 100.0%; Score 1004; DB 9; Length 1004;
Qy	961	TTAAAAAGTCAATCAGTATTCAACATCTTACATAAAAGCC 1004	XX	Best Local Similarity 100.0%; Prcd. No. 1; seq=250;
Db	961	TTAAAAAGTCAATCAGTATTCAACATCTTACATAAAAGCC 1004	XX	Indels 0; PT
Qy	15-JAN-2004	(first entry)	XX	0; Gaps 0;
ADD5519	ID	Binding domain-immunoglobulin fusion protein-associated DNA #44.	XX	CC Unidentified
ADD5519	AC	Binding domain-immunoglobulin fusion protein-associated DNA #44.	XX	Sequence 1004 BP.
ADD5519	AC	Binding domain-immunoglobulin fusion protein-associated DNA #44.	XX	PT
ADD5519	AC	Binding domain-immunoglobulin fusion protein-associated DNA #44.	XX	Query Match 100.0%; Score 1004; DB 9; Length 1004;
Qy	421	GATCTCTGACTCTGCTGCTGATGCCCTTCATCTGCTGAGCTGCTGCTGCTG 420	XX	Best Local Similarity 100.0%; Prcd. No. 1; seq=250;
Db	421	GATCTCTGACTCTGCTGCTGATGCCCTTCATCTGCTGAGCTGCTGCTGCTG 420	XX	Indels 0; PT
Qy	481	GTGCAATTGTAAGCACTTGGAGCACTATCAATGCTGAGAAATGTCAGCTG 480	XX	0; Gaps 0;
Db	481	GTGCAATTGTAAGCACTTGGAGCACTATCAATGCTGAGAAATGTCAGCTG 480	XX	CC Unidentified
Qy	541	TGTATCATTACCTTGTCAGGAGCCGCTCTTCAGGATTTCAGCTCTCC 600	XX	Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;
Db	541	TGTATCATTACCTTGTCAGGAGCCGCTCTTCAGGATTTCAGCTCTCC 600	XX	PT
Qy	541	TGTATCATTACCTTGTCAGGAGCCGCTCTTCAGGATTTCAGCTCTCC 600	XX	Query Match 100.0%; Score 1004; DB 9; Length 1004;
Db	541	TGTATCATTACCTTGTCAGGAGCCGCTCTTCAGGATTTCAGCTCTCC 600	XX	Best Local Similarity 100.0%; Prcd. No. 1; seq=250;
Qy	601	TCTTGTCCAGAACAACTGACCAAGCTTCCAGAACAACTGACCAAGCT 660	XX	Indels 0; PT

				Query	Match	Score	DB	Length
				Matches	Local Similarity	95.4%; 96.8%;	2;	1032;
				Matches	Conservative	0; Mismatches	Indels	Gaps
Db	601	TGCTTGTCCAGAACATTTGACCAAGTACTGAAAGATGTAATTACCCATAAC	660	Qy	1 CGCCAAATTAGGTTCGGTATCTCCGCTGAGCTGCTGTTCCCGCTTAGGGACC	60		
Qy	661	TTTTAACTTCAACCATGATTACTGAANITACANITGTGAAATTCTGAGGT	720	Db	1 CCCAAATTAGGTTCGGTATCTCCGCTGAGCTGCTGTTCCGGCTTAGGGACC	60		
Db	661	TTTTAACTTCAACCATGATTACTGAANITACANITGTGAAATTCTGAGGT	720	Qy	61 AGGAGAAGGGGAGCTGGAGGCTGAGCTGAGCTGCTGTTACTCGGAT	120		
Qy	721	TATTCTACTGTATTAAAGTAAATAACATAATCAAAATCTGGGGATCATGCC	780	Db	61 AGGAGAAGGGGAGCTGGAGGCTGAGCTGAGCTGCTGTTACTCGGAT	120		
Db	721	TATTCTACTGTATTAAAGTAAATAACATAATCAAAATCTGGGGATCATGCC	780	Qy	121 GTGGGCAAAGAGATGGCAAGAGCTGGAGTGTAGAAGGGGGCTGAGCCTAGG	180		
Qy	781	AGAGATTGTGGGGAGGAATGTTACAGGTTCAATTGAAATTAAATCCTAAACTT	840	Db	121 GTGGGCAAAGAGATGGCAAGAGCTGGAGTGTAGAAGGGGGCTGAGCCTAGG	180		
Db	781	AGAGATTGTGGGGAGGAATGTTACAGGTTCAATTGAAATTAAATCCTAAACTT	840	Qy	121 GTGGTGCCTAGAGATGGCAAGAGCTGGAGTGTAGAAGGGGGCTGAGCCTAGG	180		
Qy	841	ATTCTCTAGAAAATCAATAACTTGGATTTTATTCCTAAACATTTAAAAA	900	Db	121 GTGGTGCCTAGAGATGGCAAGAGCTGGAGTGTAGAAGGGGGCTGAGCCTAGG	180		
Db	841	ATTCTCTAGAAAATCAATAACTTGGATTTTATTCCTAAACATTTAAAAA	900	Qy	181 AGTGGGGCTGGAGCAAGTAAGATGGGGCAAGAGCTTCTGCCATTCTGCCAGC	240		
Qy	901	CCACTGTGAGATGTGTAATAGGFACTCTGCACTGTTGACATATAAAATA	960	Db	181 AGTGGGGCTGGAGCAAGTAAGATGGGGCAAGAGCTTCTGCCATTCTGCCAGC	240		
Db	901	CCACTGTGAGATGTGTAATAGGFACTCTGCACTGTTGACATATAAAATA	960	Qy	241 TGCTCCAAAGGGCTGGAGCTGGAGGAGCTGGAGGAGCTGGAGGAGCTGGCCTG	300		
Qy	961	TTAAAAGTCAATCAGTATTCAACATCTTACACTAAAGGCC	1044	Db	241 TGCTCCAAAGGGCTGGAGCTGGAGGAGCTGGAGGAGCTGGAGGAGCTGGCCTG	300		
Db	961	TTAAAAGTCAATCAGTATTCAACATCTTACACTAAAGGCC	1044	Qy	301 AAGAGGGCACAGCTGCTGTTCATCTCTGAGGTGTGAGCAACGGTGTGAGAGC	360		
Db				Db	301 AAGAGGGCACAGCTGCTGTTCATCTCTGAGGTGTGAGCAACGGTGTGAGAGC	360		
				Qy	361 GCAGCTCAACAGGAGMATGGAGATGGAGTGGAGTGGAGTGGCTC	420		
				Db	361 GCAGCTCAACAGGAGAATGGAGGAGTGGAGTGGCTC	420		
AC				Qy	421 GATCTCTGACCTCGTATCGGCCCTGCTGAGCTGGAGATTACAGCGAT	480		
XX				Db	421 GATCTCTGACCTCGTATCGGCCCTGCTGAGCTGGAGATTACAGCGAT	480		
DT	25-MAR-2003	(revised)		Qy	481 GTGCAATTGTGAGGACTCTGGACCACTATCAATGCTGAGAGAAATGTACCCAGA	540		
DT	22-AUG-1995	(first entry)		Db	481 CTGCAATTGTGAGGACTCTGGACCACTATCAATGCTGAGAGAAATGTACCCAGA	540		
DE	BAGE tumor rejection antigen precursor DNA.			Qy	541 TGTATCATTATCCTGTGCTGAGGCTCTGAAGATGTAATTCTACGGATAGAC	600		
XX				Db	541 TCTATCATTATCCTGTGCTGAGGCTCTCTGGATTCTACGGATAGAC	600		
OS	BAGE; tumor rejection antigen precursor; diagnosis; HLA; human leukocyte antigen MHC; major histocompatibility complex; TRAP; cancer; melanoma; ss.			Qy	601 TGTATCATTATCCTGTGCTGAGGCTCTGAAGATGTAATTCTACGGATAGAC	660		
XX				Db	601 TCTATCATTATCCTGTGCTGAGGCTCTCTGGATTCTACGGATAGAC	660		
PN	W09500159-A1.			Qy	661 TTTTAAACTTCAACCAATGTTATCTGAAAGATGTAATTCTCGTGT	720		
XX				Db	661 TTTTAAACTTCAACCAATGTTATCTGAAAGATGTAATTCTCGTGT	720		
PR	05-JAN-1995.			Qy	721 TATTCTACTGTTAAAGGTAATAATCATATCAATTAATGTTGTAATTCTCGTGT	780		
XX				Db	721 TATTCTACTGTTAAAGGTAATAATCATATCAATTAATGTTGTAATTCTCGTGT	780		
PF	10-JUN-1994;	94WO-US006534.		Qy	781 AGAGATTGTGGGGAGGG-----	812		
XX				Db	781 AGAGATTGTGGGGAGGG-----	812		
PR	17-JUN-1993;	93US-00079110.		Qy	813 GTTTCATTGAAATTAATCCTGAGCTGAAATAATCAATAAGTTGCT	840		
PR	15-FEB-1994;	94US-00196630.		Db	813 GTTTCATTGAAATTAATCCTGAGCTGAAATAATCAATAAGTTGCT	840		
PA	(LUDWIG) LUDWIG INST CANCER RES.			Qy	841 GTTTCATTGAAATTAATCCTGAGCTGAAATAATCAATAAGTTGCT	872		
XX				Db	841 GTTTCATTGAAATTAATCCTGAGCTGAAATAATCAATAAGTTGCT	900		
PI	Van Der Bruggen P, Boon-Falleur T, Coulie P, Renaud J;			Qy	873 GTTTTATCTTAAACATTTTAAACACCTGAGTGTAAATGGACTCTGTC	932		
XX				Db	901 GTTTTATCTTAAACATTTTAAACACCTGAGTGTAAATGGACTCTGTC	960		
PS	WPI: 1995-051741/07.			Qy	933 AGTATTCTGACATATACTATTAATTTAAAGTCATCACTATTCAACATCTTAA	992		
XX				Db	961 AGTATTCTGACATATACTATTAATTTAAAGTCATCACTATTCAACATCTTAA	1020		
PR	Nucleic acid coding for a tumour rejection antigen precursor - used to develop prods. for the diagnosis and therapy of cancers, partic.							
PR	melanomas.							
XX								
PS	Claim 1; Page 18; 33pp; English.							
XX								
PR	The DNA encoding the tumor rejection antigen precursor BAGE was isolated from a cDNA library prepared from RNA from human melanoma cell line M22-MEL. The sequence and its encoded protein and derived peptide may be used in the diagnosis and therapy of cancers, e.g. melanomas. (Updated on 25-MAR-2003 to correct PN field.)							
XX								
SQ	Sequence 1032 BP; 283 A; 212 C; 254 G; 283 T; 0 U; 0 Other;							

Qy	993 CACTAAGAACGCC 1004	Qy	771 GATCATGGCAGGATTTGGGAGGAAATTATCAACGGTTAATTGAAATTAAAT 830
Db	1021 CACTAAGAACGCC 1032	Db	6278 GATGTGCAGAGATTTGGGAGGAAATTATCAACGGTTAATTGAAATTAAAT 6219
<b>RESULT 6</b>			
ABU55839/c		Qy	831 CCAAAAGTATTTCCTCGAAAAATAAAATCAAATAAAGTTGATGTTTTTATCTTAAC 890
ID ABU55839 standard	DNA; 11162 BP.	Db	6218 CCAAAAGTATTTCCTCGAAAAATAAAATCAAATAAAGTTGATGTTTTTATCTTAAC 6159
XX		Qy	891 ATTAAAACCCTGAGATGTTGAAATAGGGACTGTGCAAGTATTCTGACATAC 950
AC		Db	6158 ATTAAAACCCTGAGATGTTGAAATAGGGACTGTGCAAGTATTCTGACATAC 6099
XX	15-JUL-2002 (first entry)	Qy	951 TATAAAATTAAAGTCATAGATTCAACATTTTACAAANAGCC 1004
XX	Human mutant tissue-type plasminogen activator.	Db	6098 TATAAAATTAAAGTCATAGATTCAACATTTTACAAATAAAAGCC 6045
DE		XX	
XX	Human; tissue-type plasminogen activator; mutant; thrombolytic;	XX	
KW	thrombosis; thrombus; ds.	XX	
OS	Homo sapiens.	XX	
XX	WO200220802-A1.	XX	
XX	PD 14-MAR-2002.	AC	ABL54504 standard; DNA; 11162 BP.
XX	PF 04-SEP-2000; 2000WO-CN000260.	XX	ABL54504;
XX	PR 04-SEP-2000; 2000WO-CN000260.	DT	27-MAY-2002 (first entry)
XX	(HUNA-) HUNAN ROYAL BIOTECH.	XX	Gene vector sequence SEQ ID NO 2.
XX	Xia J;	XX	Human; D chromosome; gene therapy; nucleolus; TPA; FIX;
PI		XX	KW thrombotic disease; haemophilia; vector; ds.
XX	DR WPI; 2002-292268/33.	OS	Unidentified.
XX	Construction of cell line expressing mutant human tissue-type plasminogen activator, e.g. in use for treating thrombosis by local dissolution of thrombus in blood vessels.	XX	XX WO200210375-A1.
PS	Disclosure: Page 18-23; 27pp; Chinese.	XX	XX
XX	The sequence represents the mutant human tissue-type plasminogen activator used in the invention. The invention relates to a novel cell line expressing mutant human tissue-type plasminogen activator comprising an Accession Number of CCTRC C200006. The cell line of the invention has thrombolytic activity. The method is for the construction of a cell line expressing mutant human tissue-type plasminogen activator, e.g. in use for treating thrombosis by local dissolution of thrombus in blood vessels Sequence 11162 BP; 2988 A; 2679 C; 2555 G; 2940 T; 0 U; 0 Other;	XX	XX PD 07-FEB-2002.
XX	Best Local Similarity 98.9%; Pred. No. 2.1e-125.	XX	XX PR 17-JUL-2000; 2000WO-CN000203.
Matches 528;	Conservative 6; Mismatches 0; Indels 0; Gaps 0;	XX	XX PA (XIAJ) XIA J.
Qy	471 TTACAGGGAATGTGATTTGAAAGCACTTGAGCCATATCAATGCTGTGAAAGAAA 530	XX	XX PI Xia J;
Db	6578 TTTCAGGAAATGCTGATTTTAAGCACTTGAGCCATATCAATGCTGTGAAAGAAA 6519	XX	XX DR 2002-291858/33.
Qy	531 TGTACCCAGATGTATCATATCCCTGCTGAGGGCTCTAGGATTTCAG 590	XX	XX PR 17-JUL-2000; 2000WO-CN000203.
Db	6518 TGTACCCAGATGTATCATATCCCTGCTGAGGGCTCTAGGATTTCAG 6459	XX	XX PA (XIAJ) XIA J.
Qy	591 CACATCTTCCTGCTGCTGAAACATGCAAGATGTGAAATT 710	XX	XX PT Human source chromosome originated gene leading sequence of target gene in vector for transfer into nucleolus of cell for active expression, PT useful in gene therapy and production of desirable proteins.
Db	6458 CACATCTTCCTGCTGCTGAAACATGCAAGATGTGAAATT 6399	XX	XX Disclosure: Page 63-74; 76pp; Chinese.
Qy	651 AGCGATAGACTTTAAACTCAACCAATGATTTACTGAAATTACAATGTGAAATT 710	XX	XX CC The invention relates to a human source gene leading sequence selected from DNA sequences in the short arms of the human D or G group chromosome, which does not have any important physiological function. The leading sequence is useful in gene therapy for targeting a gene in a chromosome for transfer into nucleolus of cells for active expression. The sequence may be used for the production of desirable proteins for treating, e.g. thrombotic diseases and haemophilia B with TPA and FIX, respectively. DNA sequences such as those in the short arms of the human D or G group chromosome do not have any important physiological function and are safely used in gene therapy. The present sequence is that of a gene vector sequence, useful to the invention.
Db	6398 ACGGATGAGCTTAAACTCAACCAATGATTTACTGAAATTACAATGTGAAATT 6339	XX	XX SQ Sequence 11162 BP; 2988 A; 2679 C; 2555 G; 2940 T; 0 U; 0 Other;
Qy	711 CCCTGATGTGTTATTCTACTGTATTAAGGTAATAATCATAACTGAGG 770	XX	XX SQ Query Match 52.2%; Score 524.4%; DB 6; Length 11162;
Db	6338 CCCTGATGTGTTATTCTACTGTATTAAGGTAATAATCATAACTGAGG 6279	XX	XX Best Local Similarity 98.9%; Pred. No. 2.1e-125; Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	471 TTACAGGGAATGTGATTTGAAAGCACTTGAGCCATATCAATGCTGTGAAAGAAA 530	XX	XX SQ Query Match 52.2%; Score 524.4%; DB 6; Length 11162;

Db	6578	TTCAGCAAATGCAATTGCAAGGACATTGGACCACTAAATGGTGTGAAGAGAAA	6519	CC	DNA sequences having not less than 50% homology with the above sequence, and does not have any important physiologic functions. The invention also relates to a gene vector and gene expression strategy. The leading sequence is useful in gene therapy and production of protein e.g. for treating tumour and genetic diseases. This polynucleotide sequence represents the DNA of the gene vector sequence (pGEM) of the invention
QY	531	TGTACCCAGATGATCATATTACCTTCTGCTGCTGAGGCCCTTCAAGATTCTGT	590	CC	XX
Db	6518	TGTACCCAGATGATCATATTCCCTGCTGAGGCCCTTCAAGATTCTGT	6459	CC	XX
QY	591	CACACTTCCGCTTTGTCAGAACATTAATTGACCAAGCTCTGAAAGATGTAAGTTACT	650	CC	XX
Db	6458	CACACTTCCGCTTTGTCAGAACATTAATTGACCAAGCTCTGAAAGATGTAAGTTACT	6399	SQ	Sequence 11162 BP; 2988 A; 2679 C; 2555 G; 2940 T; 0 U; 0 Other;
QY	651	ACGGATAGACTTTAACCTCAACCAATGTTACTGAAATAATACATTAATCTGAGG	710	Query	52 2%; Score 524.4; DB 6; Length 11162;
Db	6398	ACGGATAGACTTTAACCTCAACCAATGTTACTGAAATAATACATTAATCTGAGG	6339	Best Local Similarity 98.9%; Pred. No. 2.1e-125;	
QY	711	CCCTGAGTTTACTCTGTTAAAGTAAATCATTAATCTGAGG	770	Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
Db	6338	CCCTGAGTTTACTCTGTTAAAGTAAATCATTAATCTGAGG	6279	QY	471 TTACAGGATGCGATTGTAAGGACTTTGGAGCCACTATCAATGCTGTGAAGAGAAA 530
QY	771	GATCATTGCCAGAGATTGTTGGGAGAAATGTTATCAACGGTTTCAATTGAAATTAAAT	830	Db	6578 TTTCAGAATGTCGATTGTAAGGACTTTGGAGCCACTATCAATGCTGTGAAGAGAAA 6519
Db	6218	GATCATTGCCAGAGATTGTTGGGAGAAATGTTATCAACGGTTTCAATTGAAATTAAAT	6159	QY	531 TGTACCCAGATGATCATATTCCCTGCTGCTGAGGCCCTTCAAGATTCTGT
QY	831	CCAAAAGTTATTCTCGAAATAATCAATAAATGTTGCATGTTTATCTTAAAC	890	Db	6518 TGTACCCAGATGATCATATTCCCTGCTGCTGAGGCCCTTCAAGATTCTGT
Db	6218	CCAAAAGTTATTCTCGAAATAATCAATAAATGTTTCATCTTAAAC	6159	QY	591 CACACTTCCGCTTTGTCAGAACATTAATTCTGCTGAGGCCCTTCAAGATTCTGT
QY	891	ATTTAAAAACCCTGAAATGATGAAATAGGACTCTGAAATAAGCTGACATAC	950	Db	6458 CACACTTCCGCTTTGTCAGAACATTAATTCTGCTGAGGCCCTTCAAGATTCTGT
Db	6158	ATTTAAAAACCCTGAAATGATGAAATAGGACTCTGAAATAAGCTGACATAC	6099	QY	651 AGCGTAGACCTTTAACCTCAACCGATGTTACGTTACGTTACGTTACGTTAACATT
QY	951	TATATAATTATAAAAGTCAATCACTGTTCAACATCTTACACTAAAGGCC	1004	Db	6398 AGGGAGAGCTTTAACCTCAACCGATGTTACGTTACGTTACGTTACGTTAACATT
Db	6098	TATATAATTATAAAAGTCAATCACTGTTCAACATCTTACACTAAAGGCC	6045	QY	771 GATCAATTGCCAGAGATTGTTGGGGAAATCTTATCACTGTTTATCTGAAATTAAAT
RESULT 8				Db	6278 GATGTTGCCAGAGATTGTTGGGGAAATGTTATCACTGTTTATCTGAAATTAAAT
AK99658/c	ID	AAK99658 standard; DNA; 11162 BP.		QY	831 CCAAAAGTATTCTCAGAAAATCAAAATCAAAATCTGTTTATCTGAAATTAAAT
XX	AAK99658;			Db	6218 CCAAAAGTATTCTCAGAAAATCAAAATCTGTTTATCTGAAATTAAAT
AC				QY	891 ATTAAAAACCACTGAGATTGATGAAATAGGACTCTGAGATTCTGACATAC
XX				Db	6158 ATTAAAAACCACTGAGATTGATGAAATAGGACTCTGACATAC
DT	08-JUL-2002	(first entry)		QY	951 TATAAAATTAAAGTCAATCACTGTTCAACATCTTACACTAAAGGCC
XX				Db	6098 TATAAAATTAAAGTCAATCACTGTTCAACATCTTACACTAAAGGCC
DE				OS	6045
XX				PN	RESULT 9
				XX	ABA99337/c
				XX	ID ABA99337 standard; DNA; 11162 BP.
				XX	AC ABA99337;
				XX	DT 05-JUL-2002 (first entry)
				XX	XX Human TNK-TPA DNA.
				XX	XX Human; tissue-type plasminogen activator; TNK-TPA; thrombolytic;
				XX	XX thrombosis; thrombus; blood vessel; ds.
				XX	XX Homo sapiens.
				XX	XX WO200220803-A1.
				XX	XX 14-MAR-2002.
				XX	XX 16-FEB-2001; 2001WO-CN000126.
				XX	XX 17-JUL-2000; 2000WO-CN000203.
				XX	XX PA (XIAJ/) XIA J.
				XX	XX PI Xia J;
				XX	XX DR 2002-304383/34.
				XX	XX Human source chromosome-originated leading sequence as target gene in vector for expression, useful in gene therapy and production of protein e.g. for treating tumor and genetic diseases.
				XX	XX Example 2: Page 77-82; 86pp; Chinese.
				XX	XX The invention relates to a human source gene leading sequence selected from a DNA sequence in the chromosomal short arm of D and G groups, or
				XX	XX CC CC 04-SEP-2000; 2000WO-CN000260.

PA (XIAJ/) XIA J.  
 XX OS Unidentified.  
 PI XX  
 XX CN1302664 A.  
 DR XX  
 WPI; 2002-362350/39.  
 PT Cell line expressing mutant human tissue-type plasminogen activator, for  
 PT use e.g. in treating thrombosis by local dissolution of thrombus in blood  
 vessels.  
 XX  
 PS Claim 2; Page 20-25; 29pp; Chinese.  
 XX  
 CC This invention describes a novel cell line expressing mutant human tissue  
 CC -type plasminogen activator having an Accession Number of CCRCC 220006  
 CC and which has thrombolytic activity. The product of the invention can be  
 CC used for treating thrombosis by local dissolution of thrombus in blood  
 CC vessels. This sequence encodes a human tissue-type plasminogen activator  
 XX TNK-TPA described in the disclosure of the invention  
 SQ Sequence 11162 BP; 2988 A; 2679 C; 2555 G; 2940 T; 0 U; 0 Other;  
 Query Match 52.2%; Score 524.4; DB 6; Length 11162;  
 Best Local Similarity 98.9%; Pred. No. 2.1e-125;  
 Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Ps Disclosure; Page 19-25(Disclosure); 29pp; Chinese.  
 CC The present invention relates to a genetic therapy for treating  
 CC haemophilia B, which contains the gene carrier-FIX recombination body,  
 CC which uses the DNA sequence of the gene without important physiologic  
 CC function correlation on the short arm of a human chromosome in group D or  
 CC G and the DNA sequence which is 50% or more homogenous with the above  
 CC gene sequence as the pilot sequence of the therapeutic gene. This is an  
 CC oligonucleotide described in the exemplification of the invention  
 XX Sequence 13928 BP; 3744 A; 3345 C; 3126 G; 3713 T; 0 U; 0 Other;  
 SQ Query Match 52.2%; Score 524.4; DB 4;  
 Best Local Similarity 98.9%; Pred. No. 2.3e-125;  
 Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 471 TTACAGCGATGCTGCATTTCATAACACTTGGAGCCACATTATGCTGTGAAGAGAAA 530  
 Db 6578 TTTCAGCAATGCTGATTGCACTTTAAACCTTAACCTCAATGCTGTGAAGAGAAA 6519  
 Qy 531 TGTACCCAAATGTTATCATATCCTCTGTGTCGGAGCGGGCTCTTCAGGATTTCAGT 590  
 Db 651B TGTACCCAAATGTTATCATATCCTCTGTGTCGGAGCGGGCTCTTCAGGATTTCAGT 6459  
 Qy 591 CACATCTCTGCTGTTGTCAGAACACAATGTTAAAGATGTTAGTTACT 650  
 Db 645B CACATCTCTGCTGTTGTCAGAACACAATGTTAAAGATGTTACT 6399  
 Qy 651 AGCGATAGACTTTAACCTCAACAAATGTTAAATGTTGTAAT 710  
 Db 639B ACGGATAGACTTTAACCTCAACAAATGTTAAATGTTGTAAT 6339  
 Qy 711 CCCTGAGTGTATTCTACTGTGTTAAAGGTAAATAACATTAATCTGAG 770  
 Db 633B CCCTGAGTGTATTCTACTGTGTTAAAGGTAAATAACATTAATCTGAG 6279  
 Qy 771 GATCATTTGGCGAGATTGGGGAGGGAAATGTTCAACGGTTTCATGAAATTAA 830  
 Db 627B GATCGTTGGCGAGATTGGGGAGGGAAATGTTCAACGGTTTCATGAAATTAA 6219  
 Qy 831 CCAAAGATTATTCCTCGAAGAAATCAAAATCAAAATCAAAATCTAAAC 890  
 Db 621B CCAAAAGATTATTCCTCGAAGAAATCAAAATCAAAATCAAAATCTAAAC 6159  
 Qy 891 ATTTAAAAACCACTGTAAGATGTTGAAATGGGACTGTGCAATTTGACATA 950  
 Db 615B ATTTAAAAACCACTGTAAGATGTTGAAATGGGACTGTGCAATTTGACATA 6099  
 Qy 951 TATAAAATTAAAGTCATCAGTTCAACATCTTTACATGAAATTAAAGCC 1.004  
 Db 6098 TATAAAATTAAAGTCATCAGTTCAACATCTTTACATGAAATTAAAGCC 6045  
 DB 04-DEC-2001 (first entry)  
 XX Haemophilia B genetic therapy related oligonucleotide #3.  
 DE Haemophilia B, coagulant; ds.  
 KW Haemophilia B; coagulant; ds.

RESULT 10  
 AAH77500/C  
 ID AAH77500 standard; DNA; 13928 BP.  
 XX  
 AC AAH77500;  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Haemophilia B genetic therapy related oligonucleotide #3.  
 XX  
 KW Haemophilia B; coagulant; ds.

831 CCAAAAGTTATTCCTCGAAGAAATCAAAATCAAAATGTTCACTGTTTATTCTAAAC 890  
 8984 CCAAAAGTTATTCCTCGAAGAAATCAAAATCAAAATGTTTATTCTAAAC 8925  
 891 ATTAAAAAACCTGTAAGATGTTGAAATGGGACTGTGCAATTTGACATA 950  
 8924 ATTAAAAAACCTGTAAGATGTTGAAATGGGACTGTGCAATTTGACATA 8865  
 951 TATAAAATTAAAGTCATCAGTTCAACATCTTTACATGAAATTAAAGCC 1.004

Db	8864 TATAAAATTAAAAAGTCATCGATTCAACATCCTTACACTAAAAGCC 8811 RESULT 11 AAK9657/C ID AAK9657 standard; DNA; 107602 BP. XX AC AAK9657; XX DT 08-JUL-2002 (first entry) DE DNA of the PAC clone sequence. XX CYOstatic; human; gene leading sequence; short arm D Group; tumour; XX KW gene therapy; short arm G group; genetic disease; PAC clone sequence; ds. XX Homo sapiens.	Qy 771 GATCATGCCAGAGATTGTTGGGAGGAATGTATCACAGGTTTCACTAAATTAAAT 830 Db 77448 GATCGTGCAGAGATTGTTGGGAGGAATGTATCACAGGTTTCACTAAATTAAAT 77389 Qy 831 CCAAAGTATTCTCAGAAATCAATAAGGTGATGTTTATTCCTAAAC 890 Db 77388 CCAAAGTATTCTCAGAAATCAATAAGGTGATGTTTATTCCTAAAC 77329 Qy 891 ATTTAAACCACTGAGATGATTAAGGACTGCTGAGATTCTGACATAC 950 Db 77328 ATTTAAACCACTGAGATGATTAAGGACTGCTGAGATTCTGACATAC 77269 Qy 951 TATAAAATTAAAAAGTCATCGATTCAACTCTTACACTAAAGGCC 1004 Db 77268 TATAAAATTAAAAAGTCATCGATTCAACTCTTACACTAAAGGCC 77215
		RESULT 12 ABL54503/C ID ABL54503 standard; DNA; 107612 BP. XX AC ABL54503; XX DT 27-MAY-2002 (first entry) XX Human PAC clone sequence SEQ ID NO 1.
		XX DB Human; D chromosome; G chromosome; gene therapy; nucleolus; TPA; FIX; XX KW thrombotic disease; haemophilia; PAC clone; ds. XX PI Homo sapiens. XX PN WO200210375-A1. XX PR XX XX PS 16-FEB-2001; 2001WO-CN000126. XX PR 17-JUL-2000; 2000WO-CN000203. XX PA (XIAJ) XIA J. XX PI Xia J; XX DR 2002-304383/34. XX PT Human source chromosome originated leading sequence as target gene in PT vector for expression, useful in gene therapy and production of protein PT e.g. for treating tumor and generic diseases. XX PS Example 1: Page 28-77; 86pp; Chinese. XX CC The invention relates to a human source gene leading sequence selected CC from a DNA sequence in the human short arm of D and G groups, or CC DNA sequences having not less than 50% homology with the above sequence, CC and does not have any important physiological functions. The invention CC also relates to a gene vector and gene expression strategy. The leading CC sequence is useful in gene therapy and production of protein e.g. for CC treating tumour and genetic diseases. This polynucleotide sequence CC represents the DNA of the PAC clone sequence of the invention XX SQ Sequence 107502 BP; 36346 A; 19466 C; 18991 G; 32779 T; 0 U; 20 Other; Query Match 52.2%; Score 524.4%; DB 6; Length 107602; Best Local Similarity 98.9%; Pred. No. 5..8e-125; Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
		XX CC The invention relates to a human source gene leading sequence selected CC from DNA sequences in the short arms of the human D or G group CC chromosomes, which does not have any important physiological function. The CC leading sequence is useful in gene therapy for targeting a gene in a CC vector for transfer into nucleolus of cells for active expression. The CC sequence may be used for the production of desirable proteins for CC treating, e.g. thrombotic diseases and haemophilia B with TPA and FIX, CC respectively. DNA sequences such as those in the short arms of the human CC D or G group chromosome do not have any important physiological function CC and are safely used in gene therapy. The present sequence is that of a CC human PAC clone sequence, useful to the invention XX SQ Sequence 107612 BP; 36302 A; 19500 C; 19011 G; 32769 T; 0 U; 30 Other;
		Query Match 52.2%; Score 524.4%; DB 6; Length 107612; Best Local Similarity 98.9%; Pred. No. 5..8e-125; Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
		Qy 471 TTACAGCGATGTCGATTGTGCACTTCAATGCTGTTGAGGAGGCTCTTCACTAAATGTTAAGGAA 530 Db 77748 TTTCGCAATGTCGATTGTGCACTTCAATGCTGTTGAGGAGGCTCTTCACTAAATGTTAAGGAA 77689 Qy 531 TGATCCCCAGATGTCATCAATTCCTTGCTGCGAGGAGGCTCTTCACTAAATGTTAAGGAA 590 Db 77688 TTGACCCATGTCATCAATTCCTTGCTGCGAGGAGGCTCTTCACTAAATGTTAAGGAA 77629 Qy 591 CACATCTCTGCTGTCAGACACATTGACCAAGTCTCTGAAATAAGTAACTTCAATCTGAGG 770 Db 77628 CACATCTCTGCTGTCAGACACATTGACCAAGTCTCTGAAATAAGTAACTTCAATCTGAGG 77569 Qy 651 ACGATAACGTTAACATCAACATTAACTGAAATGTTAAGTGTAAATT 710 Db 77568 ACGATAACGTTAACATCAACATTAACTGAAATGTTAAGTGTAAATT 77509 Qy 711 CCCTGAGGTAACTCTGTATTAAGTAACTATCAATCTGAGG 770 Db 77508 CCCTGAGGTAACTCTGTATTAAGTAACTATCAATCTGAGG 77449
		Db 77749 TTTCGCAATGTCGATTGTGCACTTCAATGCTGTTGAGGAGGCTCTTCACTAAATGTTAAGGAA 77690

CC claims the KGF PCR primers AAC87589 and AAC87590 for use in detecting all CC or part of the KGF gene. The nucleic acid sequences can be used to detect CC abnormalities in chromosomal region 9p11 and thus give an indication of CC an individual's risk of developing a 9p11-associated condition. The present sequence represents human chromosomal region 9q11.

XX Sequence 57728 BP; 16960 A; 12164 C; 11741 G; 16863 T; 0 U; 0 Other;

XX Query Match 18.1%; Score 181.8; DB 4; Length 57728;

XX Best Local Similarity 89.9%; Pred. No. 4.e-36;

XX Matches 195; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

XX SQ

QY 591. CACATCTCTGCTTGTCAAGAACATGCAAGCTCTCGAAAGATGTAAGTTACT 650

Db 77629 CACATCTCTGCTTGTCAAGAACATGCAAGCTCTCGAAAGATGTAAGTTACT 77570

QY 651. ACCGATAGCTCTTAAACTCACCATTACGTTACGAAATACTGGTTAAATT 710

Db 77569 ACCGATAGCTCTTAAACTCACCATTACGTTACGAAATACTGGTTAAATT 77510

QY 711. CCTCTGAGGTTATTCTACTGTATTAAAGGTTAATACTATAATTAATCTGAGG 770

Db 77509 CCTCTGAGGTTATTCTACTGTATTAAAGGTTAATACTATAATTAATCTGAGG 77450

QY 777. GATCATGCCAGAGATGTTGGGGAGGAAATGTTACAACGGTTACATGAAATAAT 830

Db 77449 GATGTTGCCAGAGATGTTGGGGAGGAAATGTTACAACGGTTACATGAAATAAT 77390

QY 831. CCAAAAAGTTATTCTCTGAAAAAATCAAATAAGTTGCATGTTTTTATCTAAAC 890

Db 77389 CCAAAAAGTTATTCTCTGAAAAAATCAAATAAGTTGCATGTTTTTATCTAAAC 77330

QY 891. ATTTCAAAACCCTGTGAAATGTTGAAATAGGACTGTGCAAGTATTTCGACATAC 950

Db 77329 ATTTCAAAACCCTGTGAAATGTTGAAATAGGACTGTGCAAGTATTTCGACATAC 77270

QY 951. TATAAAATTATTAAAAGTCATCACTATTCACATTTACATTTACATTTACATCAAC 1004

Db 77269 TATAAAATTATTAAAAGTCATCAATGTTACATTTACATTTACATTTACATCAAC 77216

RESULT 14

ID AAS44735 standard; DNA; 6292 BP.

XX ID AAS44735 standard; DNA; 6292 BP.

XX AC AAS44735;

XX DT 18-DEC-2001 (first entry)

XX DE Tumour suppressor gene derived chemically modified sequence #459.

XX KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;

XX DT tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;

XX KW cytosine methylation; ds.

XX OS Homo sapiens.

XX PN WO200168912-A2.

XX PD 20-SEP-2001.

XX PF 15-MAR-2001; 2001WO-EP002955.

XX JP20000287684-A.

XX PR 15-MAR-2000; 2000DE-01013847.

XX PR 06-APR-2000; 2000DE-01019058.

XX PR 07-APR-2000; 2000DE-01019173.

XX PR 30-JUN-2000; 2000DE-0103529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX PA Olek A, Piepenbrock C, Berlin K;

XX PI XX

XX DR WPI; 2001-602752/68.

XX PT Fragments of chemically modified genes associated with tumor suppressor genes and oncogenes, useful in designing primers and probes for analyzing diseases associated with cytosine methylation state e.g. cancer.

XX PS Claim 1; SEQ ID NO 459; 27pp; English.

XX CC The invention relates to a nucleic acid comprising a sequence of 18 bases of a segment of chemically prereated DNA (CP DNA) e.g. with bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and

XX CC Abnormalities in this region of the short arm of chromosome 9 is thought to be associated with miscarriage and cancer, as an ovarian cancer patient with a history of miscarriage was found to have a chromosomal inversion inv(9) (p11;q13). The 9p11 region contains the gene encoding keratinocyte growth factor (KGF), and the invention also specifically

CC 500 are missing from the sequence listing) sequences (Ss) and sequences complementary to (Se). The nucleic acid may be a peptide nucleic acid-CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of CC nucleotide polymorphisms and also to be used in an array for analysing CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The CC probes can also be used in a method for ascertaining genetic and/or epigenetic parameters for the diagnosis of existing CC diseases or the predisposition to specific diseases, by analysing CC cytosine methylations. The parameters may be compared to another set of CC genetic and/or epigenetic parameters, the differences serving as basis CC for diagnosis and/or prognosis events which are disadvantageous to patients. The present sequence is one of the 533 genomic sequences derived from tumour suppressor genes and oncogenes. Note: The sequence data for this patent did not form part of the printed specification, but CC was obtained in electronic format directly from WIPO at ftp://ipo.int/pub/published\_pct\_sequences

XX Sequence 6292 BP; 1736 A; 99 C; 1304 G; 3153 T; 0 U; 0 Other;  
 SQ Query Match 14.9%; Score 149.8; DB 4; Length 6292;  
 Best Local Similarity 80.6%; Pred. No. 3 3e-28;  
 Matches 175; Conservative 0; N mismatches 42; Indels 0; Gaps 0;  
 Db 5001 CGTTATTTAGGGTCTCGGTATCTCCCGTGAAGCTGTCGTCAGGAGGCC 60  
 QY 1 CGCCCAATTAGGGTCTCGGTATCTCCCGTGAAGCTGTCGTCAGGAGGCC 60  
 Db 5001 CGTTATTTAGGGTCTCGGTATCTCCCGTGAAGCTGTCGTCAGGAGGCC 60  
 QY 61 AGGAGAAAGGGAGGTGAGGCTGGAGCTGGACCTGTAACACCTGGTCTGTTCTACTCTGTGTG 120  
 Db 5061 AGGAGAAAGGGAGGTGAGGCTGGACCTGTAACACCTGGTCTGTTCTACTCTGTGTG 120  
 QY 121 GTGGTGGCACAGAATGGAGGCTGGAGCTGGAGGTAGGAGGGCTGACCGTAGG 180  
 Db 5121 GTGGTGGCACAGAATGGAGGCTGGAGCTGGAGGTAGGAGGGCTGACCGTAGG 180  
 QY 181 AGTGGGGCTGGAGGTAAATGGGGCAAGAGGGT 217  
 Db 5181 AGTGGGGCTGGAGGTAAATGGGGCAAGAGGGT 217

RESULT 15  
 AAC19397  
 ID AAC19397 standard; CDNA; 251 BP.  
 XX  
 AC AAC19397;  
 DT 06-OCT-2000 (first entry)  
 XX Human secreted protein 5, EST, SEQ ID NO: 23472.  
 XX Human; 5' EST; expressed sequence tag; secreted protein; CDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 XX Homo sapiens.  
 OS  
 PN EP1033401-A2.  
 XX  
 PD 06-SEP-2000.  
 XX 21-FEB-2000; 2000085-002000610.  
 XX 26-FEB-1999; 990US-0122487P.  
 PA (GEST ) GENSET.  
 XX Dumas Milne Edwards J, Ductier A, Giordano J;

XX WPI; 2000-500381/45.  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PR diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 23472; 71pp + Sequence Listing; English.  
 PS The present sequence is one of a large number of 5' ESTs derived from CC mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dt primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors  
 XX Sequence 251 BP; 49 A; 50 C; 99 G; 50 T; 0 U; 3 Other;  
 SQ Query Match 10.6%; Score 106.2; DB 3; Length 251;  
 Best Local Similarity 88.8%; Pred. No. 1.6e-17;  
 Matches 111; Conservative 3; Mismatches 11; Indels 0; Gaps 0;  
 QY 261 AAGGAGAAAGTCCCCTGTTGAGCTGGAGGTGGCCTGAAGAGCGCAAGCTCTGTGC 320  
 Db 1 AAGGAGAAAGTCCCCTGTTGAGCTGGAGGTGGCCTGAAGAGCGCAAGCTCTGTGA 60  
 QY 321 TTATCTCTTGTAGGTGTGGAGCAAGGTATGGAGACGGTCAACGGGAAATA 380  
 Db 61 TTATCTCTTGTGGGTGTGGAGGTGGCTGAGGACTCAGGTCACAGGAGTGTA 120  
 QY 381 GGAGG 385  
 Db 121 GGAGG 125  
 Search completed: July 9, 2004, 07:09:19  
 Job time : 641 secs

This Page Blank (uspto)

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: July 9, 2004, 05:43:15 ; Search time 105 Seconds  
(without alignments)  
5306.388 Million cell updates/sec

Title: US-10-081-108-1

Perfect score: 1004  
Sequence: CGCCAAATTAGGGTCTCGG.....ATCTTTACACTAAAAAGGC 1004

Scoring table: IDENTITY NUC Gap 10.0 , Gapext 1.0

Searched: 682709 seqs, 27475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgcn2\_6/.ptodata/2/ina/SA\_COMB.seq:  
2: /cgcn2\_6/.ptodata/2/ina/5B\_COMB.seq:  
3: /cgcn2\_6/.ptodata/2/ina/6A\_COMB.seq:  
4: /cgcn2\_6/.ptodata/2/ina/6B\_COMB.seq:  
5: /cgcn2\_6/.ptodata/2/ina/BCUTUS\_COMB.seq:  
6: /cgcn2\_6/.ptodata/2/ina/backfile1.seq:  
7: /cgcn2\_6/.ptodata/2/ina/backfile2.seq:  
8: /cgcn2\_6/.ptodata/2/ina/backfile3.seq:  
9: /cgcn2\_6/.ptodata/2/ina/backfile4.seq:  
10: /cgcn2\_6/.ptodata/2/ina/backfile5.seq:  
11: /cgcn2\_6/.ptodata/2/ina/backfile6.seq:  
12: /cgcn2\_6/.ptodata/2/ina/backfile7.seq:  
13: /cgcn2\_6/.ptodata/2/ina/backfile8.seq:  
14: /cgcn2\_6/.ptodata/2/ina/backfile9.seq:  
15: /cgcn2\_6/.ptodata/2/ina/backfile10.seq:  
16: /cgcn2\_6/.ptodata/2/ina/backfile11.seq:  
17: /cgcn2\_6/.ptodata/2/ina/backfile12.seq:  
18: /cgcn2\_6/.ptodata/2/ina/backfile13.seq:  
19: /cgcn2\_6/.ptodata/2/ina/backfile14.seq:  
20: /cgcn2\_6/.ptodata/2/ina/backfile15.seq:  
21: /cgcn2\_6/.ptodata/2/ina/backfile16.seq:  
22: /cgcn2\_6/.ptodata/2/ina/backfile17.seq:  
23: /cgcn2\_6/.ptodata/2/ina/backfile18.seq:  
24: /cgcn2\_6/.ptodata/2/ina/backfile19.seq:  
25: /cgcn2\_6/.ptodata/2/ina/backfile20.seq:  
26: /cgcn2\_6/.ptodata/2/ina/backfile21.seq:  
27: /cgcn2\_6/.ptodata/2/ina/backfile22.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description	
1	1004	100.0	1004	2	US-08-389-360-1	Sequence 1, Appli	
2	1004	100.0	1004	3	US-09-038-328-1	Sequence 1, Appli	
3	1004	100.0	1004	4	US-09-435-524-10	Sequence 20, Appli	
4	1004	100.0	1004	4	US-09-1382-497-1	Sequence 1, Appli	
5	1004	100.0	1004	4	US-08-079-110A-1	Sequence 1, Appli	
6	96.6	96.2	1032	1	US-08-196-630A-1	Sequence 1, Appli	
7	96.6	96.2	1032	3	US-08-573-186-1	Sequence 1, Appli	
8	96.6	96.2	1032	3	US-10-027-983-11	Sequence 11, Appli	
9	83	8.3	39200	4	US-09-535-008-18	Sequence 58, Appli	
10	82.4	8.2	859	4	US-09-439-261-34	Sequence 34, Appli	
11	82.4	8.2	4698	4	US-09-227-613-33	Sequence 33, Appli	
12	82.4	8.2	4698	4	US-09-844-634-17	Sequence 17, Appli	
c	13	82.4	8.2	15602	4	US-08-305-6918-19	Sequence 19, Appli
c	14	82.4	8.2	35100	1	PCT-US93-06251-19	Sequence 10, Appli
c	15	82.4	8.2	35100	5	Sequence 19, Appli	Sequence 10, Appli
c	16	82.4	8.2	39982	4	US-09-820-924-3	Sequence 3, Appli
c	17	82.4	8.2	84495	4	US-09-797-906-3	Sequence 3, Appli
c	18	82.4	8.2	319608	4	US-09-539-333D-1	Sequence 1, Appli
c	19	82.4	8.2	319608	4	US-09-679-409-1	Sequence 1, Appli
c	20	82.2	8.2	7210	2	US-08-257-963B-10	Sequence 10, Appli
c	21	82.2	8.2	7210	4	US-08-367-841A-10	Sequence 10, Appli
c	22	82.2	8.2	7210	5	PCT-US95-07201-10	Sequence 10, Appli
c	23	82.2	8.2	14581	4	US-08-520-373D-4	Sequence 4, Appli
c	24	82.2	8.2	22481	4	US-08-367-841A-43	Sequence 43, Appli
c	25	82.2	8.2	22481	5	PCT-US95-07201-43	Sequence 2, Appli
c	26	82.2	8.2	22484	4	US-09-875-223-2	Sequence 2, Appli
c	27	82.2	8.2	22484	4	US-09-875-114-2	Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-08-389-360-1  
Sequence 1, Application US/08389360  
Patent No. 587017  
GENERAL INFORMATION:  
APPLICANT: van der Bruggen et al.  
TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/389,360  
FILING DATE: Herewith  
CLASSIFICATION: 436  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/196,630  
FILING DATE: February 15, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/079,110  
FILING DATE: June 17, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REFERENCE/DOCKET NUMBER: LUD 5310.2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3894  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1004  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-389-360-1

Query Match Score 1004; DB 2; Length 1004;  
Best Local Similarity 100.0%; Pred. No. 3.7e-258;  
Matches 1004; Conservative 0; Missmatches 0; Indexes 0; Gaps 0;

Patent No. 6110694  
 GENERAL INFORMATION:  
 APPLICANT: van der Bruggen et al.  
 TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES NUMBER OF SEQUENCE: 9  
 TITLE OF INVENTION: THEREOF  
 CORRESPONDENCE ADDRESS:  
 STREET: 805 Third Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/038,328  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/389,360  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/079,110  
 FILING DATE: June 17, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,994  
 REFERENCE/DOCKET NUMBER: LUD 5310.2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEX/FAX: (212) 638-3884  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1004;  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-038-328-1

Query Match 100.0%; Score 1004; DB 3; Length 1004;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-120;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCATTAGGGTCTCCGATATCCCGTCTGGCTGAGGAC 60  
 Db 1 CGCCATTAGGGTCTCCGATCTCCGCTGAGGAC 60  
 Qy 61 AGGAAGCCGGAGCTGGGTGAGCTGTACACGGTGTGAG 120  
 Db 61 AGGAAGCCGGAGCTGGGTGAGCTGTACACGGTGTGAG 120  
 Qy 121 GTGGTGCACAGAGTGGCTGAGGTTTACGGTGTGAG 180  
 Db 121 GTGGTGCACAGAGTGGCTGAGGTTTACGGTGTGAG 180  
 Qy 121 AGAGATTGTCGGAGGAATAGTATCAACGTTTACATGAA 840  
 Db 781 AGAGATTGTCGGAGGAATAGTATCAACGTTTACATGAA 840  
 Qy 721 TATTCTACTGTTAATAAGGTAAATAATCATTAATCTGAGGT 780  
 Db 721 TATTCTACTGTTAATAAGGTAAATAATCATTAATCTGAGGT 780  
 Qy 781 AGAGATTGTCGGAGGAATAGTATCAACGTTTACATGAA 840  
 Db 781 AGAGATTGTCGGAGGAATAGTATCAACGTTTACATGAA 840  
 Qy 841 ATTCCTCAGAAAATCAATAAAGTTGCAATTTAACATTAA 900  
 Db 841 ATTCCTCAGAAAATCAATAAAGTTGCAATTTAACATTAA 900  
 Qy 901 CCACTAGTGAATGATGTAATAGGAACTGTCAGATTTGCA 960  
 Db 901 CCACTAGTGAATGATGTAATAGGAACTGTCAGATTTGCA 960  
 Qy 961 TTAAAAAGTCATCATTCACATTTACACTAAAGCC 1004  
 Db 961 TTAAAAAGTCATCATTCACATTTACACTAAAGCC 1004  
 Qy 301 AGAGCCCAAGCTCTGCTTCACTTCAGGTTGTCAGGAGG 360  
 Db 301 AGAGCCCAAGCTCTGCTTCACTTCAGGTTGTCAGGAGG 360  
 Qy 361 GCAGCTAACAGGAGAACATGAGGATAGGAGGATGGCTC 420  
 Db 361 Sequence 1, Application US/09038328-1  
 RESULT 2  
 US-09-038-328-1

Db 361 GCAGCTAACGGCCAATAGGAGAGATTGAGGTTCAGTGTCAAGGATGGTCTC 420  
 Qy 421 GATCCTCTGAACTCTGATCGCCCTTCAAAGTGCCGATTACGCCAT 480  
 Db 421 GATCCTCTGAACTCTGATCGCCCTTCAAAGTGCCGATTACGCCAT 480  
 Qy 481 GTGCATTTTAAAGCACATTGGAGCACAATCAATGCTGTGAAGAAAATGACCC2GA 540  
 Db 481 GTGCATTTTAAAGCACATTGGAGCACAATCAATGCTGTGAAGAAAATGACCC2GA 540  
 Qy 541 TGTATCATTTACCTTCTGCTGAGGCCGCTCTTCAAGGTTTCAATCTCC 600  
 Db 541 TGTATCATTTACCTTCTGCTGAGGCCGCTCTTCAAGGTTTCAATCTCC 600  
 Qy 601 TGCTTGTCCAGAAACATGCAAGCTGAGGAGCTGAAAGTAGTTACCATAC 660  
 Db 601 TGCTTGTCCAGAAACATGCAAGCTGAGGAGCTGAAAGTAGTTACCATAC 660  
 Qy 661 TTAAACATCAACAAATGATTAATCTGATAAATTCCTGAGNGT 720  
 Db 661 TTAAACATCAACAAATGATTAATCTGATAAATTCCTGAGNGT 720  
 Qy 721 TATTCTACTGTATAAACGTAATAATCATATAATCTGGGGATCATGCC 780  
 Db 721 TATTCTACTGTATAAACGTAATAATCATATAATCTGGGGATCATGCC 780  
 Qy 781 AGAGATTCTGGGAGGAAATCTTACAGGTTTCAATGAAATTATCCAAAGTT 840  
 Db 781 AGAGATTCTGGGAGGAAATCTTACAGGTTTCAATGAAATTATCCAAAGTT 840  
 Qy 841 ATTCTCTCAGAAAATCAATAAAGTTGATGTTTCAATGTTTCAATGTTAA 900  
 Db 841 ATTCTCTCAGAAAATCAATAAAGTTGATGTTTCAATGTTTCAATGTTAA 900  
 Qy 901 CCACTGAAATGATGAAATAGGACTCTGCAAGTATTCTGCATATAAAATA 960  
 Db 901 CCACTGAAATGATGAAATAGGACTCTGCAAGTATTCTGCATATAAAATA 960  
 Qy 961 TTAAAAAGTCATCAATGATCAACATCTTACACTAAAGCC 1004  
 Db 961 TTAAAAAGTCATCAATGATCAACATCTTACACTAAAGCC 1004  
 RESULT 3  
 US-09-056-105-20  
 Sequence 20, Application US/09056105  
 ; Patent No. 6287569  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIPPS, THOMAS J.  
 ; APPLICANT: WU, YUNQI  
 ; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRABELLULAR  
 ; TITLE OF INVENTION: PROCESSING  
 ; FILE REFERENCE: 233/221  
 ; CURRENT APPLICATION NUMBER: US/09/056,105  
 ; EARLIER APPLICATION NUMBER: 60/043,467  
 ; EARLIER FILING DATE: 1997-04-10  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: FastSEQ For Windows Version 3.0  
 ; SEQ ID NO 20  
 ; LENGTH: 1004  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-056-105-20

Query Match 100.0%; Score 1004; DB 3; Length 1004;  
 Best Local Similarity 100.0%; Pred. No. 3.e-58;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCAAATTAGGCTCTCCCGTATCTCCCGTAGCTGAGCTAGAGCC 60  
 Db 1 CGCCAAATTAGGCTCTCCCGTATCTCCCGTAGCTGAGCTAGAGCC 60

RESULT 4  
 US-09-435-524-1  
 ; Sequence 1, Application US/09435524  
 ; Patent No. 6465184  
 ; GENERAL INFORMATION:  
 ; APPLICANT: van der Bruggen et al.  
 ; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES









```

; RESULT 9
; Sequence 11, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; ATTORNEY: Mark P. Roach
; TITLE OF INVENTION: RANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 392000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; OTHER INFORMATION: unknown
; LOCATION: 137740
; NAME/KEY: misc feature
; OTHER INFORMATION: unknown
; LOCATION: 138122 ..(138221)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 145507
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 151957
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: 1542066
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 154217
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: 1542066 ..(154216)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 174657 ..(174756)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 186224 ..(186323)
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: 195242 ..(195341)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 202703
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: 202771 ..(202870)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 206246 ..(215603)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 218126 ..(218225)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 227487 ..(227536)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 230157 ..(230256)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 222717 ..(222846)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 223981 ..(224030)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 227487 ..(227536)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 232599 ..(232398)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 236552 ..(236651)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 238789 ..(248788)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: exon
; LOCATION: 118288 ..(119101)
; OTHER INFORMATION: exon 1C
; NAME/KEY: exon:inttron junction
; LOCATION: 151129 ..(151130)
; OTHER INFORMATION: exon 5:inttron 5
; NAME/KEY: exon:inttron junction
; LOCATION: 239248 ..(239249)
; OTHER INFORMATION: exon 9:inttron 9
; NAME/KEY: exon:inttron junction
; LOCATION: 348578 ..(348579)
; OTHER INFORMATION: exon 10:inttron 10
; NAME/KEY: inttron
; LOCATION: 348579 ..(381838)
; OTHER INFORMATION: intron 10
; NAME/KEY: inttron:exon junction
; LOCATION: 386185 ..(386186)
; OTHER INFORMATION: inttron 11:exon 12
; US-10-027,983-11

Query Match 386 AGATGGAGTTCACTGTTCAAGCAGGATGGCTCGATCTCTGACCTCGTGAATCCGCC 445
Best Local Similarity 89.9%; Pred. No. 2,4e-11;
Matches 89; Conservative 0; Mismatches 10; Gaps 0;

Db 252283 AGACGGAGTTCACTGTTAGCAGGATGGCTCGATCTCTGACCTCGTGAATCCGCC 2523429
Qy 446 GCCTTGCGCTTCCAAGGCGGAGATAACAGGGATGNGC 484
Db 252343 GCCTCAGCTTCCAAGGCTGGATACAGGGATGAGC 252381

; RESULT 10
; Sequence 58, Application US/09535008
; Patent No. 6465629
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavriqian, Sean V.
; APPLICANT: Teng, David H.-F.
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; TITLE OF INVENTION: AND OTHER CANCER TYPES
; FILE REFERENCE: 23118-259
; CURRENT APPLICATION NUMBER: US/09/535,008
; CURRENT FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: U.S. 60/125,806
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 7.0
;
```

SEQ ID NO: 58  
*i LENGTH: 859*  
*i TYPE: DNA*  
*i ORGANISM: Homo sapiens*  
US-09-535-008-58

Query Match 8.2%; Score 82.4; DB 4; Length 859;  
Best Local Similarity 89.0%; Pred. No. 2e-12; Indels 0; Gaps 0;  
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 385 GAGATGGAGTTCACTGTTCAAGATGGTCACTCGACCTCGTGTGATCGCC 444  
Db 80 GAGACGGGGTTTCACCGTTAACCGTTAGCCAGATGGTTCGATCTCCGACCTCGTGTGATCGCC 139

Qy 445 CGCCATTGGCTTCGAAGGCGGATACAGGGATGTGC 484  
Db 140 CGCCATTGGCTTCGAAGGCGGATACGGGATTCAGGGTGTGCC 179

RESULT 11  
US-09-433-261-34  
*i Sequence 34, Application US/094339261*  
*i Patent No. 6428930*  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Mukerji, Pradip  
APPLICANT: Leonardi, Amanda E.  
APPLICANT: Huang, Yung-Sheng  
TITLE OF INVENTION: HUMAN DESATURASE GENE AND USES THEREOF  
FILE REFERENCE: 6295 US P2  
CURRENT APPLICATION NUMBER: US/09/439,261  
PRIORITY FILING DATE: 1999-11-12  
PRIOR APPLICATION NUMBER: US 08/833,610  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: PCT/US98/07422  
PRIOR FILING DATE: 1998-04-10  
PRIOR APPLICATION NUMBER: US 09/227,613  
PRIOR FILING DATE: 1999-01-08  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO: 34  
LENGTH: 4698  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-433-261-34

Query Match 8.2%; Score 82.4%; DB 4; Length 4698;  
Best Local Similarity 89.0%; Pred. No. 4.e-12; Indels 0; Gaps 0;  
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 385 GAGATGGAGTTCACTGTTCAAGGATGGTCACTCGACCTCGTGTGATCGCC 444  
Db 3324 GAGACGGGGTTCACTGTTCACTGTTAGCCAGATGGTCTCGACTCGTGTGATCGCC 3383

RESULT 12  
US-09-227-613-33  
*i Sequence 33, Application US/09227613A*  
*i Patent No. 6432884*  
GENERAL INFORMATION:  
APPLICANT: MUKERJI, Pradip  
APPLICANT: LEONARD, Amanda E.  
APPLICANT: HUANG, Yung-Sheng  
TITLE OF INVENTION: HUMAN DESATURASE GENE AND USES THEREOF  
FILE REFERENCE: 6295 US P1  
CURRENT APPLICATION NUMBER: US/09/227,613A  
PRIOR APPLICATION NUMBER: US 08/833,610  
PRIOR FILING DATE: 1997-04-11

RESULT 13  
US-09-844-634-17/C  
*i Sequence 17, Application US/09844634*  
*i Patent No. 6410324*  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF TUMOR NECROSIS FACTOR RECEPTOR 2 EXPRESS  
FILE REFERENCE: RTS-016  
CURRENT FILING DATE: 2001-04-27  
NUMBER OF SEQ ID NOS: 174  
SEQ ID NO: 17  
LENGTH: 15602  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-844-634-17

Query Match 8.2%; Score 82.4%; DB 4; Length 15602;  
Best Local Similarity 89.0%; Pred. No. 7.e-12;  
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 385 GAGATGGAGTTCACTGTTCAAGGATGGTCTCGATCTCCNGACCTCGTGTGATCGCC 444  
Db 1945 GAGATGGGGTTCACTGTTAGCCAGATGGTCTCGATCTCCGACCTCGTGTGATCGCC 1886

Qy 445 CGCCATTGGCTTCGAAGGCGGATACAGGGATGTGC 484  
Db 1885 CCCCTGGCTTCGAAGGCTGGATTACAGGGTGTGAGC 1846

RESULT 14  
US-08-306-691B-19  
*i Sequence 19, Application US/08306691B*  
*i Patent No. 5734039*  
GENERAL INFORMATION:  
APPLICANT: Calabretta, Bruno  
APPLICANT: Storski, Tomasz  
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
CORRESPONDENCE ADDRESS:  
ADRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.  
STREET: Two Penn Center, Suite 1800  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB  
COMPUTER: IBM PS/2

```

OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 35100 base Pairs
TYPE: nucleic acid
STRANDBNESS: double
TOPOLOGY: linear
US-08-306-691B-19

Query Match 8.2%; Score 82.4; DB 1; Length 35100;
Best Local Similarity 89.0%; Prod. No. 1..le-11;
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
PCT-US93-06251-19

Query Match 8.2%; Score 82.4; DB 5; Length 35100;
Best Local Similarity 89.0%; Prod. No. 1..le-11;
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
PCT-US93-06251-19

Qy 385 GAGATGGAACTTCACTGTTCACTGTTCACTGATGGCTCGATCTGACTCTGATCGGCC 444
Db 28167 GAGATGGGTTCACCATTTAGCCGATGGCTCGATCTGACTCTGATCGGCC 28226

Qy 445 CGCCTTGGCCTTCCAAGTGCCGAGATTACAGCGATGTC 484
Db 28227 CACCTGGCTCCAAAGTGTGAGATTACAGGCATGAGC 28266

SEQUENCE: 19
Search completed: July 9, 2004, 10:09:37
Job time : 110 secs

```

```

RESULT 15
PCT-US93-06251-19
Sequence 19, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wicksstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/TUS93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DiGiilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELEPHONE: 516-742-4243
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 35100 base Pairs

```

Copyright GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 08:57:39 ; Search time 694 Seconds  
 (without alignments)  
 7050.978 Million cell updates/sec

Title: US-10-081-108-1  
 Perfect score: 1004  
 Sequence: 1 CGCCAATTAGGGCTTCGGG.....ATCTTACACAAAAAGCC 1004

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 3183909 seeds, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0  
 Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Published Applications NA:  
 1: /cgm2/\_/protoData/\_/pupnra/us07\_PUBCOMB.seq:\*

Database :  
 1: /cgm2/\_/protoData/\_/pupnra/us07\_PUBCOMB.seq:\*

## ALIGNMENTS

RESULT 1  
 US-10-207-655-80  
 Sequence 80, Application US/10/207655  
 Publication No. US2003118552A1  
 GENERAL INFORMATION:  
 APPLICANT: Ledbetter, Jeffrey A.  
 APPLICANT: Hayden-Ledbetter, Martha S.  
 TITLE OF INVENTION: IMMUNOGLOBULIN FUSION PROTEINS  
 FILE REFERENCE: 390069\_401C1  
 CURRENT APPLICATION NUMBER: US/10/207,655  
 CURRENT FILING DATE: 2002-07-25  
 NUMBER OF SEQ ID NCS: 426  
 SOFTWARE: Patentin version 3.0  
 SEQ ID NO: 80  
 /cgn2\_6\_ptodata1/pubnpa/us08\_pub.seq;  
 /cgn2\_6\_ptodata1/pubnpa/us08\_pubcomb.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us09a\_pubcomb.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us09b\_pubcomb.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us09c\_pubcomb.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us09\_new\_pub.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us09\_new\_pub.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us10a\_pubcomb.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us10b\_pubcomb.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us10c\_pubcomb.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us10\_new\_pub.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us60\_new\_pub.seq;\*  
 /cgn2\_6\_ptodata1/pubnpa/us60\_pubcomb.seq;\*

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

תְּהִלָּה

Result No.	Score	Query			Length	DB ID	Description	
		Match	Length	DB				
1	1004	100.0	1004	15	US-10-207-655-80		Sequence 80, Appl	
2	1004	100.0	1004	15	US-10-081-108-1		Sequence 1, Appl	
c	3	524.4	52.2	1162	17	US-10-063-404-1		Sequence 1, Appl
c	4	524.4	52.2	13958	13	US-10-362-916-1		Sequence 1, Appl
c	5	330.4	32.9	524	15	US-10-029-386-1510		Sequence 1510, Ap
c	6	330.4	32.9	524	15	US-10-029-388-8714		Sequence 8714, Ap
c	7	164.6	16.4	345	15	US-10-029-389-1839		Sequence 18339, A
c	8	164.6	16.4	538	15	US-10-029-388-4583		Sequence 4583, Ap
c	9	164.6	16.4	538	15	US-10-029-388-2249		Sequence 22449, Ap
c	10	149.8	14.9	172	15	US-10-029-316-15211		Sequence 15211, A
c	11	149.8	14.9	153	15	US-10-029-316-15211		Sequence 461, Appl
c	12	93.6	9.3	6292	13	US-10-021-714-461		Sequence 1252, Ap
c	13	88.4	8.8	31516	13	US-10-087-112-1252		Sequence 128006, Ap
c	14	88.4	8.8	407	13	US-10-027-632-128026		Sequence 128006, Ap
c	15	88.4	8.8	12	12	US-10-027-632-128026		Sequence 128006, Ap

Query	Match	100.0%	Score 1004;	DB 15;	Length 1004;
	Best Local Similarity	100.0%;	Pred. No. 2.	8e-247;	
	Matches 1004;	Conservative	0;	Mismatches	0;
				Indels	Gaps
Qy	1	CGCAATTAGGGTTCGGTATCCGGCTGAGCTCTGGTCCCGCTTAGAGAAC	60		
Db	1	CGCAATTAGGGTTCGGTATCCGGCTGAGCTCTGGTCCCGCTTAGAGAAC	60		
Qy	61	AGGAAAGGGAGTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	120		
Db	61	AGGAAAGGGAGTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	120		
Qy	121	GTGTGGCAACAGAGATGGCAGCCAGCTGGACTTAGGAGGGCTGAGGGTAG	180		
Db	121	GTGTGGCAACAGAGATGGCAGCCAGCTGGAGTTAGGAGGGCTGAGGGTAG	180		
Qy	181	AGTGGGGCTGGAGGAGTAAATGGCGGCCAGGGGTTTCTGCATTGCTGCCAG	240		

MEDIUM: TYPE: 3.5 inch 1.44 Mb storage diskette  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US 10/081,108  
 FILING DATE: 20-Feb-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/435,524  
 FILING DATE: 08-NO- US 2003/0138854A1-1999  
 APPLICATION NUMBER: 09/038,328  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: 08/079,110  
 FILING DATE: June 17, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,894  
 REFERENCE DOCKET NUMBER: LUD 5310.2  
 TELECOMMUNICATION INFORMATION:  
 Matches 1004; Pred. No. 2, 8e-24; Indels 0; Gaps 0;  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-081-108-1

Query 1 CGCCCAATTAGGGCTCCGGTATCCTCGCTGAGCTCCGGCTTAGAGGACCC 60  
 Best Local Similarity 100.0%; Pred. No. 2, 8e-24;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 481 GTGCATTGTAAGGAACTTGGACCATATCAAAATGCTGTGAAGAGAAATGTTACCCAGA 540  
 Db 481 GTGCATTGTAAGGAACTTGGACCATATCAAAATGCTGTGAAGAGAAATTACCCAGA 540

Query 541 TGTATCATATCCCTGTGATTCGCCGCCCTTGCGCTTCAAAATGCGGATTACCGGAT 480  
 Db 541 TGTATCATATCCCTGTGATTCGCCGCCCTTGCGCTTCAAAATGCGGATTACCGGAT 480

Query 481 GTGCATTGTAAGGAACTTGGACCATATCAAAATGCTGTGAAGAGAAATGTTACCCAGA 540  
 Db 481 GTGCATTGTAAGGAACTTGGACCATATCAAAATGCTGTGAAGAGAAATTACCCAGA 540

Query 541 TGTATCATATCCCTGTGATTCGCCGCCCTTGCGCTTCAAAATGCGGATTACCGGAT 480  
 Db 541 TGTATCATATCCCTGTGATTCGCCGCCCTTGCGCTTCAAAATGCTGTGAATTCCTGAGTC 600

Query 541 TGTATCATATCCCTGTGATTCGCCGCCCTTGCGCTTCAAAATGCTGTGAATTCCTGAGTC 600

Db 541 TGTATCATATCCCTGTGATTCGCCGCCCTTGCGCTTCAAAATGCTGTGAATTCCTGAGTC 600

Query 601 TGCTTGTCAAGAACATGACCAAGCTCTGAAGATGTAAGTTACTAGCATAGAC 660  
 Db 601 TGCTTGTCAAGAACATGACCAAGCTCTGAAGATGTAAGTTACTAGCATAGAC 660

Query 661 TTAAAACCTCAACCAATGTTACTGAAATAACAAATGTTGAAATTCCTGAGTC 720  
 Db 661 TTAAAACCTCAACCAATGTTACTGAAATAACAAATGTTGAAATTCCTGAGTC 720

Query 721 TATTCTACTGTATAAAAGCTTAATAATCATATACTGAGGGATCATTCGC 780  
 Db 721 TATTCTACTGTATAAAAGCTTAATAATCATATACTGAGGGATCATTCGC 780

Query 781 AGAGATTGTTGGAGGGAAATGTTCATGAAATAATCCAAAAAGTT 840  
 Db 781 AGAGATTGTTGGAGGGAAATGTTCATGAAATAATCCAAAAAGTT 840

Query 841 ATTCCTCAAAATTCAATAALGTTTCATGTTTAAACATTAAATAAA 900  
 Db 841 ATTCCTCAAAATTCAATAAACTGTTCATGTTTAAACATTAAATAAA 900

Query 901 CCACGTGAGATGTTGAAATAGGACTGTGCACTTGTGACATAATAAAATAA 960  
 Db 901 CCACGTGAGATGTTGAAATAGGACTGTGCACTTGTGACATAATAAAATAA 960

Query 961 TTAAAAGTCATGATGTTGAAATAGGACTGTGCACTTGTGACATAAAAGCC 1004  
 Db 961 TTAAAAGTCATGATGTTGAAATAGGACTGTGCACTTGTGACATAAAAGCC 1004

Query 301 AAAGACGCCACAGCTCTGCTTCATCTTCTGAGGTGTCGCCACGGTGTGAGGACCC 360  
 Db 301 AAAGACGCCACAGCTCTGCTTCATCTTCTGAGGTGTCGCCACGGTGTGAGGACCC 360

Query 361 GCAGCTCAAACGGACAAATGGAGGAAGTGGAGTTCACTGTCAGGAGATGTC 420  
 Db 361 GCAGCTCAAACGGACAAATGGAGGAAGTGGAGTTCACTGTCAGGAGATGTC 420

Query 421 GATCTCTGACCTCGTCAATGGAGTTGAGGAAATGGAGTTGAGGATTCAGGGAT 480  
 Db 421 GATCTCTGACCTCGTCAATGGAGTTGAGGAAATGGAGTTGAGGATTCAGGGAT 480

RESULT 2  
 US-10-081-108-1  
 Sequence 1, Application US/10081108  
 Publication No. US2003/0138854A1  
 GENERAL INFORMATION:  
 APPLICANT: van der Bruggen et al.  
 TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
 WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
 THEREOF  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Feife & Lynch  
 STREET: 805 Third Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022  
 COMPUTER READABLE FORM:



RESULT 5  
 US-10-029-386-1510/c  
 ; Sequence 1510, Application US/10029386  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS TWO  
 ; FILE REFERENCE: AEONICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029,386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288  
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
 SEQ ID NO: 8714  
 LENGTH: 524  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AF254983.1  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95  
 OTHER INFORMATION: EST HUMAN HIT: AI138404.1, EVALUATION 0.00e+00  
 OTHER INFORMATION: NT HIT: AU163201.2, EVALUATION 0.00e+00  
 OTHER INFORMATION: SWISSPROT HIT: P55200, EVALUATION 9.00e-04  
 US-10-029-386-8714

Query Match 32.9%; Score 330.4; DB 15; Length 524;  
 Best Local Similarity 98.2%; Pred. No. 2.e-74;  
 Matches 334; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 530  
 Db 185 TTTCAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 244  
 Qy 531 TGTACCCAGATGTATATTATCCTGTGTGCAGAGGCCGCTTCAAGTTTCACT 590  
 Db 245 TGTACCCAGATGTATATTATCCTGTGTGCAGAGGCCGACACTTCAGATTCTAGT 304  
 Qy 591 CACATOTTCCTGCCTGCTGTCAGACATGGCAAGTGTAACTGAGGTAGTTACT 650  
 Db 305 CACATTCCTGCCTGTCAGACATGGCAAGTGTAACTGAGGTAGTTACT 364  
 Qy 651 ACGCATAGACTTTAACACTCAACCAATGTTAATGTTAATTGTCAATT 710  
 Db 365 ACGCATAGACTTTAACACTCAACCAATGTTAATGTTAATTGTCAATT 424  
 Qy 711 CCCTGATGTGTATTACTGTATTAAAGTTAATCATATAATCTGAGG 770  
 Db 425 CCCTGATGTGTATTACTGTATTAAAGTTAATCATATAATCTGAGG 484

RESULT 6  
 US-10-029-386-8714  
 ; Sequence 8714, Application US/10029386  
 ; Publication No. US20030194704A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS TWO  
 ; FILE REFERENCE: AEONICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029,386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288

Query Match 32.9%; Score 330.4; DB 15; Length 524;  
 Best Local Similarity 98.2%; Pred. No. 2.e-74;  
 Matches 334; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 530  
 Db 185 TTTCAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 244  
 Qy 531 TGTACCCAGATGTATATTATCCTGTGTGCAGAGGCCGCTTCAAGTTTCACT 590  
 Db 245 TGTACCCAGATGTATATTATCCTGTGTGCAGAGGCCGACACTTCAGATTCTAGT 304  
 Qy 591 CACATOTTCCTGCCTGCTGTCAGACATGGCAAGTGTAACTGAGGTAGTTACT 650  
 Db 305 CACATTCCTGCCTGTCAGACATGGCAAGTGTAACTGAGGTAGTTACT 364  
 Qy 651 ACGCATAGACTTTAACACTCAACCAATGTTAATGTTAATTGTCAATT 710  
 Db 365 ACGCATAGACTTTAACACTCAACCAATGTTAATGTTAATTGTCAATT 424  
 Qy 711 CCCTGATGTGTATTACTGTATTAAAGTTAATCATATAATCTGAGG 770  
 Db 425 CCCTGATGTGTATTACTGTATTAAAGTTAATCATATAATCTGAGG 484

RESULT 7  
 US-10-029-386-18339/c  
 ; Sequence 18339, Application US/10029386  
 ; Publication No. US20030194704A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS TWO  
 ; FILE REFERENCE: AEONICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029,386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288

Query Match 32.9%; Score 330.4; DB 15; Length 524;  
 Best Local Similarity 98.2%; Pred. No. 2.e-74;  
 Matches 334; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 530  
 Db 340 TTTCAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 281  
 Qy 531 TGTACCCAGATGTATATTATCCTGTGTGCAGAGGCCGCTTCAAGTTTCACT 590  
 Db 280 TGTACCCAGATGTATATTATCCTGTGTGCAGAGGCCGCTTCAAGTTTCACT 221  
 Qy 591 CACATCTTCCTGCTGCTGTCAGACATGGCAAGTGTAACTGAGGTAGTTACT 650  
 Db 220 CACATTCCTGCCTGCTGTCAGACATGGCAAGTGTAACTGAGGTAGTTACT 161  
 Qy 651 ACGCATAGACTTTAACACTCAACCAATGTTAATGTTAATTGTCAATT 710  
 Db 160 ACGCATAGACTTTAACACTCAACCAATGTTAATGTTAATTGTCAATT 101  
 Qy 711 CCCTGATGTGTATTACTGTATTAAAGTTAATCATATAATCTGAGG 770  
 Db 100 CCCTGATGTGTATTACTGTATTAAAGTTAATCATATAATCTGAGG 41  
 Qy 771 GATCATGGCCAGAGATTGTGGGAGGAATGTGTATCAA 810  
 Db 40 GATCATGGCCAGAGATTGTGGGAGGAATGTGTATCAA 524

SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
SEQ ID NO: 18339  
LENGTH: 345  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049849.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7  
OTHER INFORMATION: NT HIT: 91:161:65974, EVALUE 0.00e+00  
OTHER INFORMATION: EST HUMAN HIT: BE890156.1, EVALUE 7.00e-65  
OTHER INFORMATION: SWISSPROT HIT: Q13072, EVALUE 2.00e-06  
US-10-029-386-18339

Query Match Score 164.6; DB 15; Length 345;  
Best Local Similarity 97.7%; Pred. No. 6.7e-32; Indels 0; Gaps 0;  
Matches 167; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 215 GGTTTTCTGGCATGTCGCCAACCTGCTCCAAAGCAAGCTGATGAGGAGTCCTC 274  
Db 282 GGTTTTCTGGCATGTCGCCAACCTGCTCCAAAGCAAGCTGATGAGGAGTCCTC 223  
QY 275 TTGTGTGAGCTGGAGTTGAGCTTGAGCTGAGAGGCAAGTCTGTCTCATCTCTGG 334  
Db 222 TTGTGTGAGCTGGAGTTGAGCTTGAGCTGAGAGGCAAGTCTGTCTCATCTCTGG 163  
QY 335 TTGTGGAGCCACGTGTATGGAGACGGAGCTCAACAGGAGCATAGGAGG 385  
Db 162 TTGTGGAGCCACGTGTATGGAGACGGAGCTCAACAGGAGCATAGGAGG 112

RESULT B  
US-10-029-386-4583/c  
Sequence 4533, Application US/10029386  
Publication No. US20030194704A1  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
SEQ ID NO: 4583  
LENGTH: 538  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049849.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7  
OTHER INFORMATION: NT HIT: 91:1476:3849, EVALUE 4.00e-06  
OTHER INFORMATION: EST HUMAN HIT: BE890156.1, EVALUE 1.00e-64  
OTHER INFORMATION: EST\_HUMAN HIT: BE890156.1, EVALUE 1.00e-64  
US-10-029-386-4583

Query Match Score 164.6; DB 15; Length 538;  
Best Local Similarity 97.7%; Pred. No. 9e-32; Indels 0; Gaps 0;  
Matches 167; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 215 GGTTTTCTGGCATGTCGCCAACCTGCTCCAAAGCAAGCTGATGAGGAGTCCTC 274

Db 366 GTTTTCTGGATTTGCTGCCAACCTGCTCCAAAGCAAGCTGATGAGGAGTCCTC 307  
Qy 275 TTGTGTGAGCTGGAGTTGAGCTGAGCTGAGAGGCAAGTCTGTCTCATCTCTGG 334  
Db 306 TTGTGTGAGCTGGAGTTGAGCTGAGCTGAGAGGCAAGTCTGTCTCATCTCTGG 247  
Qy 335 TTGTGGAGCCACGTGTATGGAGACGGAGCTCAACAGGAGCATAGGAGG 385  
Db 246 TTGTGGAGCCACGTGTATGGAGACGGAGCTCAACAGGAGCATAGGAGG 196

RESULT 9  
US-10-029-386-22449  
Sequence 22449, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
SEQ ID NO: 22449  
LENGTH: 172  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AF254983.1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95  
OTHER INFORMATION: SWISSPROT HIT: P55200, EVALUE 2.00e-04  
EST HUMAN HIT: AI138404.1, EVALUE 6.00e-82  
NT HIT: AL163201.2, EVALUE 3.00e-92  
US-10-029-386-22449

Query Match Score 164.2; DB 15; Length 172;  
Best Local Similarity 98.2%; Pred. No. 5.5e-32;  
Matches 166; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 471 TTACAGGATGNGCATTTGTAAGGACTTTGGAGCACTATCAATGCTGTGAAGAGAAA 530  
Db 4 TTACAGGATGNGCATTTGTAAGGACTTTGGAGCACTATCAATGCTGTGAAGAGAAA 63  
Qy 531 TGTACCCAGATGATCATATTCTCTGTCAGGATTCAGT 590  
Db 64 TGTACCCAGATGATCATATTCTCTGTCAGGATTCAGT 123

Qy 591 CACATCTTCCCTGCTTGTCCAGAAACATTGACCAAGCTCTGTGAAGAGAT 639  
Db 124 CACATCTTCCCTGCTTGTCCAGAAACATTGACCAAGCTCTGTGAAGAGAT 172

RESULT 10  
US-10-029-386-15211/c  
Sequence 15211, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288



```

; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 128026
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-128026

Query Match          8 .8%; Score 88.4; DB 13; Length 407;
Best Local Similarity 98 .9%; Pred. No. 2.8e-12;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   471 TTACAGCGATGCGATTGTTAAGCACTTGGGCCATTACATGCTGTGAGAGAA 530
Db   90 TTTCAAGCGATGCGATTGTTAAGCACTTGGGCCATTACATGCTGTGAGAGAA 31

Qy   531 TGTACCCAGATGATCATATCCTGTGCT 560
Db   30 .TGTACCCAGATGATCATATCCTGTGCT 1

RESULT 14
US-10-027-632-128027/C
Sequence 128027, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 128028
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-128028

```

```

Query Match          8 .8%; Score 88.4; DB 13; Length 407;
Best Local Similarity 98 .9%; Pred. No. 2.8e-12;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   471 TTACAGCGATGCGATTGTTAAGCACTTGGGCCATTACATGCTGTGCT 560
Db   90 TTTCAAGCGATGCGATTGTTAAGCACTTGGGCCATTACATGCTGTGCT 1

Qy   531 TGTACCCAGATGATCATATCCTGTGCT 560
Db   30 .TGTACCCAGATGATCATATCCTGTGCT 1

Search completed: July 9, 2004, 12:21:50
Job time : 697 secs

```

```

Query Match          8 .8%; Score 88.4; DB 13; Length 407;
Best Local Similarity 98 .9%; Pred. No. 2.8e-12;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy   471 TTACAGCGATGCGATTGTTAAGCACTTGGGCCATTACATGCTGTGCT 530
Db   90 TTTCAAGCGATGCGATTGTTAAGCACTTGGGCCATTACATGCTGTGCT 31

```

This Page Blank (uspto)

This Page Blank (uspto)



Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krywinski, M.I., Skalicka, U., Smailus, D.E., Schein, A., Schein, J.E., Jones, S.J. and Marra, M.A.	Db	325 AAAGCCGCGACAGCTCTGGCTTCATCTTGAGGTGGCACCGTGTAGGAAGCG 384
TITLE	Qy	361 GCAGCTAAAGGAGCAATAGGAG-----GATGGAGTTCACT 400
JOURNAL	Db	385 GCAGCTAAAGGAGCAATAGGAGCAAGTGTGAAATGGACTTCAGT 444
PUBMED	Qy	401 GTGTGAGCCGAGATGTCGATCCTAACCTGATGGTGGCTCAA 460
REFERENCE	Db	445 GTGTGAGCCGAGATGTCGATCCTAACCTGATGGTGGCTCAA 504
AUTHORS	Qy	461 AGTGCCTGAGATA-----CAGCGATGTG 483
JOURNAL	Db	505 ATGGCGGAGTACAGGTTATGAGAGATTCTCCCTTACCGATG 564
Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	Qy	543 CATTGTAAAGCTTGGGCCATTAATGCTGAAAGAAATGCACTG 543
NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	Db	565 CATTGTAAAGCTTGGGCCATTAATGCTGAAAGAAATGCACTG 624
COMMENT	Qy	544 ATCATATCCTGGCTGAGGCGCCCTTACAGATTGACATTTCTGC 603
Tissue Procurement: ARCC/DCRD/DTP	Db	625 ATCATATCCTGGCTGAGGCGCCCTTACAGATTGACATTTCTGC 684
cDNA Library Preparation: Life Technologies, Inc.	Qy	604 TTTGTCAGACACATGTGCAAGATGTAAGTTACTACGCCATAGCTT 663
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	Db	685 TTTCGCGAGACAGCTGACCAGTGTGCAAGATGTAAGTTACTAGGATAGCTT 744
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305	Qy	664 TAAACTCAACCAATGTTACTGAAATAACAAATGTTAAATCCCTGAGTGTAT 723
Web site: <a href="http://www.smgc.stanford.edu">http://www.smgc.stanford.edu</a>	Db	745 TAAACTCAACCAATGTTACTGAAATAACAAATGTTAAATCCCTGAGTGTAT 804
Contact: (Dickson, Mark) mcd@paxil.stanford.edu	Qy	724 TCTACTGTGTTAAAGCTAAATCATCTAAATCATGAGGTCAATTGCGACA 783
Dickson, M., Schmutz, J., Grimmood, J., Rodriguez, A., and Myers, R. M.	Db	805 TCTACTGTGTTAAAGGTATAATACTAAATCATGAGGTCAATTGCGACA 864
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL site: <a href="http://image.llnl.gov">http://image.llnl.gov</a>	Qy	784 GATTGTGGGAGGGAATGTTACAACGSGTTTCATGAAATTCAAAGTATT 843
Series: IRAK Plate: 22 Row: o Column: 15	Db	865 GATTGTGGGAGGGAATGTTACAACGSGTTTCATGAAATTCAAAGTATT 924
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557346	Qy	844 TCCTCGAAAMATCAATAAGTTTCATGTTTATCTAACTTTAAACCA 903
This clone has the following problem: frame shifted.	Db	925 TCCTCGAAAMATCAATAAGTTTCATGTTTATCTAACTTTAAACCA 984
Location/Qualifiers	Qy	904 CTGTAGATGTGTAATAAGGACTGTGAGTATTGACATATACTATAATTATA 963
1.	Db	985 CTGTAGATGTGTAATAAGGACTGTGAGTATTGACATATACTATAATTATA 1044
.organism="Homo sapiens"	Qy	964 AAAAGCTAACATGATTCACATCTTACACTAAAAGCC 1004
mol type="mRNA"	Db	1045 AAAAGCTAACATGATTCACATCTTACACTAAAAGCC 1085
obj_xref="taxon:3606"		
clone="IMAGE_39164151"		
tissue_type="Skin, melanotic melanoma."		
clone.Tib="NH MGIC_72"		
/lab_host="DH10B"		
note="vector: pCMV-SPORT6"		
FEATURES	source	RESULT 2
source	Qy	CA423957/C
Match 92.0%; Score 923.8; DB 11; Length 2037;	Db	732 bp mRNA linear EST 07-NOV-2002
Best Local Similarity 94.4%; Pred. No. 3.6e-220; N mismatches 0; Indels 57; Gaps 2;	Qy	DEFINITION UI-H-FE1.bed-p-04-0-UI s1 NCI CGAP FE1 Homo sapiens cDNA clone
Matches 1002; Conservative	Db	ACCESSION CA423957
Qy	25 CGCCCAATTAGGTCTCCGGATCCTCGGTGAGCTGCTCACTGGATG	VERSION CA423957.1 GI:24786683
Db	61 AGGAGAGGGGAGCTGGAGCTGAGCTGCTAAACCCGGCTGAGCTG	SOURCE Homo sapiens (human)
Db	85 AGGAGAGGGGAGCTGGAGCTGAGCTGCTAAACCCGGCTGAGCTG	ORGANISM Homo sapiens
Qy	121 GTGGTGGCAAAGAGATGGCAGCTGGAGCTGGAGCTGGAGCTGGAG	REFERENCE NCI-CGAP
Db	145 GTGGTGGCAAAGAGATGGCAGCTGGAGCTGGAGCTGGAGCTGGAG	AUTHORS http://www.ncbi.nlm.nih.gov/cnicgap.
Db	181 AGTGGGGCTGGAGCACTAAGATGGGGCGCAGAGCTGGCTGCTG	TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Unpublished (1997)
Db	205 ACTGGGGCTGGAGCACTAAGATGGGGCGCAGAGCTGGCTGCTG	JOURNAL Contact: Robert Strausberg, Ph.D.
Qy	241 TGCTCCAAGCAGGCTGATGAGGGAGTGGCTGAGCTGGAGCTGGAG	COMMENT Email: cgaps-r@mail.nih.gov.
Db	265 TGCTCCAAGCAGGCTGATGAGGGAGTGGCTGAGCTGGAGCTGGAG	Tissue Procurement: James Martin
Qy	301 AAGACCGCACAGCTCTGGCTTCATCTTGAGGTGAGGTGAGGAG	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
		cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares bento-soares@uiowa.edu  
 The following repetitive elements were found in this cDNA sequence: 391-493, >ALU 493-590, >ALU  
 Seq\_Primer: M13 FORWARD  
 POLYA=Yes  
 Location/Qualifiers  
 source 1. .732  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref=taxon:9006"  
 /clone=JUL-H-FE1-ped-P-04-0-UI"  
 /tissue\_type="Cell lines"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP\_FEL"  
 /note="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP\_FEL is a normalized cDNA library derived from a pool of mRNA obtained from 3 cell lines from grade II chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first strand cDNA contains a library tag (dR18 tail). The sequence tag for this library is CGCTAACGGAC. The cell lines were provided by Dr James Martin from the University of Iowa.  
 AG TISSUE=Human grade 2 chondrosarcoma cell line pool TAG LIB=JUL-H-FE1  
 TAG SEQ=CGCTAACGGAC

## ORIGIN

Query Match 44.9%; Score 450.8; DB 14; Length 732;  
 Best Local Similarity 81.5%; Pred. No. 8.1e-102; Mismatches 0; Indels 108; Gaps 2;  
 Matches 595; Conservative

QY	243	CTCCAGGCCAGGCGATGAGGAGACTCCCTGGTAGCTGGAGGTGGCCGAA	302
DB	732	CTCCAGGCCAGGCGATGAGGAGACTCCCTGGTAGCTGGAGGTGGCCGAA	673
QY	303	GACGGCAAGCTCTGTCATCTCTGAGGTGGAGCCACGGGTGAGAGCGC	362
DB	672	GACGGCAAGCTCTGTCATCTCTGAGGTGGAGCCACGGGTGAGAGCGC	613
QY	363	AGCTCAAGAGGATAAGGAGATGGATTCACTTGAGCCAGGATGGCTCGA	422
DB	612	AGCTCAAGAGGATAAGGAGATGGATTCACTTGAGCCAGGATGGCTCGA	553
QY	423	TCTCTGACCT--GTGATCGCCGCTGGCTTCCAAA-----	461
DB	552	GCTCTGGTCAAGTGATGCCGCCCTAGGCTCCAAAAGTGCTGGATTATGGCG	493
QY	462	-----	461
DB	492	CTGGAGTGCATGGACGATCTGGCTCACTGGCACGGCATCCCTGGCTGGTCAA	433
QY	462	-----	461
DB	432	GCGATTCTCCGCCCTCAACCTCCCGAGTAGCTGGATACAGGATGTGCAATTGGAG	373
QY	495	CACPTGGCCACTATCAAATGCTGTGAGAAATGACCGAGATATTAATCT	554
DB	372	CACPTGGCCACTATCAAATGCTGTGAGAAATGACCGAGATATTAATCT	313
QY	555	TGTGCTGAGAGGCCGCTCTTCAGGATTCACTCCTCCGTCTGAGAGAA	614
DB	312	TGTGCTGAGAGGCCGCTCTTCAGGATTCACTCCTCCGTCTGAGAGAA	253

QY 615 CACATTGACCAAGCTCCCTGAAAGATGTAAGTTTAAACTTCAC 674  
 Db 252 CACATTGACCAAGCTCCCTGAAAGATGTAAGTTTAAACTTCAC 193  
 QY 675 CAATGTTATTACTGAAATAACAAAGCTGAAATTCTACTGTTATTCTACTGTAT 734  
 Db 192 CAATGTTATTACTGAAATAACAAAGCTGAAATTCTACTGTTATTCTACTGTAT 133  
 QY 735 TAAAGGTTATAATACATAATCATATAATCTGGGATCATGGCAGAGATTGGGG 794  
 Db 132 TAAAGGTTATAATACATAATCATATAATCTGGGATCATGGCAGAGATTGGGG 73  
 QY 795 AGGGAAATGTTATCAACGGTTCATGAAATAATCCTCAGAAAA 854  
 Db 72 AGGGAAATGTTATCAACGGTTCATGAAATAATCCTCAGAAAA 13  
 RESULT 3  
 LOCUS BE890756 699 bp mRNA linear EST 20-OCT-2000  
 DEFINITION 60143191P1 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:3916415 5', mRNA sequence.  
 ACCESSION BE890756  
 VERSION BE890756.1 GI:10349397  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi; Mgcidae; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 699)  
 AUTHORS NIH\_MGC\_HTC://mgc.ncbi.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-r@mail.nih.gov  
 Tissue Procurement: ARCC/DCD/PDR  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov/plate:11AM9741> row: e column: 24  
 High quality sequence stop: 407.  
 Location/Qualifiers 1. .699  
 FEATURES /organism="Homo sapiens"  
 source /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3916415"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NTH MGC\_72"  
 /note="Organ: Skin; Vector: PCMV-SPORT6; Site\_1: Not I; Site\_2: Sall; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."  
 ORIGIN

Query Match 34.8%; Score 349.6; DB 10; Length 699;  
 Best Local Similarity 89.1%; Pred. No. 1.8e-76;  
 Matches 454; Conservative 0; Mismatches 30; Gaps 6;  
 QY 1 CGCCAAATTAGCTCTCGGTATCTCCGCTGAGCTGCTGTCGGCTTAGGGACC 60  
 Db 4 CGCCAAATTAGCTCTCGGTATCTCCGCTGAGCTGCTGTCGGCTTAGGGACC 63  
 QY 61 AGGAGAGGGGAGGGCTGGAGGCTGCTGAACTCGCTGACTGGATG 120



/clone="KAT00969"  
 /clone\_lib="Sugano Homo sapiens cDNA library"

**ORIGIN**

Query Match Score 245.8; DB 9; Length 300;  
 Best Local Similarity 93.5%; Pred. No. 1.4e-50;  
 Matches 267; Conservative 0; Mismatches 17; Indels 1; Gaps 0;

Qy 17 CCGPATCTCCGCTGAGCNGCTCTGTCGGCTTAGGACCAAGGAGGAGCT 76  
 Db 1 CCAGATCCCTCGAGCTGGTCCGAGTAGAACAGAGAAGGGAGCT 60

Qy 17 CGAGCTGGAGCTGTAAACACGGTCTCGTGTCACTCTGGATGGCTGCACAGAGA 136  
 Db 61 GGAGCTGGAGCTGTAAACACATGGCTGTGGTGTGGAGTGGAGAGA 120

Qy 17 CCTGGATCTCCGCTAGCTCTGTCGGCTTAAGGACCAAGGAGATT GGAGCT 75  
 Db 137 TGCGAGGCCAGGTGGAGTGTTAGAAGGGCCCTGAGGGTGGCGGTAGGAGTGGCAG 196

Qy 77 GGAGGTGGACCTGCTTAACACCGTGGCTCTCTACTCTGGATGTTGGCACAGGGA 136  
 Db 121 TGCGAGGCCGCTGTAACTGGAGCTGTGGAGCTGGAGGCTGAGCAG 180

Db 76 TGAGGGTGGACCTGCTTAACACCGTGGCTCTCTGGATGTTGGCACAGGGA 135  
 Qy 197 TAAGATGGGGCCAGGGGT 217  
 Db 181 TAGGAGGGGGCTGGAGCAGT 201

**RESULT 7**

A1138404 LOCUS 345 bp mRNA linear EST 28-OCT-1998  
 DEFINITION qd84b05.x1 Scores test is NHT Homo sapiens cDNA clone IMAGE:1736145  
 3, similar to TR:014686 ALR.; mRNA sequence.

Qy 196 TAAGATGGGGCCAGGGGTTCCTGGCATGGCTGGAGGAGCT 256  
 Db 196 TAAGATGGGGCCAGGGGTTCCTGGCATGGCTGGAGGAGCT 255

Qy 257 GATGAGGAGGAGTCCCCTGTGGAGCTGGAGGTGGACCTGA 301  
 Db 256 GATGAGGAGGAGTCCCCTGTGGAGCTGGCTGGAGCTGGACCTGA 300

**RESULT 6**

BM449877 LOCUS 940 bp mRNA linear EST 05-FEB-2002  
 DEFINITION AGENCOURT\_6393056 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5528612  
 5', mRNA sequence.

Qy 258 GATGAGGAGGAGTCCCCTGTGGAGCTGGAGGTGGACCTGA 301  
 Db 258 GATGAGGAGGAGTCCCCTGTGGAGCTGGCTGGAGCTGGACCTGA 300

**RESULT 5**

BM449877 LOCUS 940 bp mRNA linear EST 05-FEB-2002  
 DEFINITION AGENCOURT\_6393056 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5528612  
 5', mRNA sequence.

Qy 259 GATGAGGAGGAGTCCCCTGTGGAGCTGGAGGTGGACCTGA 301  
 Db 259 GATGAGGAGGAGTCCCCTGTGGAGCTGGCTGGAGCTGGACCTGA 300

**ORGANISM**

Homo sapiens (human)

**FEATURES**

source

organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5528612"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NTH MGC\_72"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primers: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

**ORIGIN**

Query Match 15.8%; Score 158.4; DB 9; Length 345;  
 Best Local Similarity 99.4%; Pred. No. 1.1e-28;  
 Matches 159; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**ORGANISM**

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE**

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

**AUTHORS**

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

**TITLE**

Tumor Gene Index

**JOURNAL**

Unpublished (1997)

**COMMENT**

Contact: Robert Strausberg, Ph.D.  
 Email: cgaps@nlm.nih.gov  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLM ac:  
 www-bio.lnl.gov/bbrp/image/image.html

**VERSION**

BM449877.1 GI:18498917

**KEYWORDS**

EST.

**SOURCE**

Homo sapiens (human)

**ORGANISM**

Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homo.

**REFERENCE**

NIH-MGC http://mgc.nci.nih.gov/

**AUTHORS**

National Institutes of Health, Mammalian Gene Collection (MGC)

**TITLE**

Unpublished (1999)

**JOURNAL**

Contact: Robert Strausberg, Ph.D.  
 Email: cgaps@nlm.nih.gov

**FEATURES**

source

organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1736145"  
 /sex="male"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares testis\_NHT"  
 /note="Vector: pCR3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5'-TGTACCCATCTGAGTGGAGGGCCCCAATTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pR73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Facima Bonaldo."

**ORIGIN**

Query Match 15.8%; Score 158.4; DB 9; Length 345;  
 Best Local Similarity 99.4%; Pred. No. 1.1e-28;  
 Matches 159; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 8  
**BX282448/c** BX282448 385 bp mRNA linear EST 04-MAR-2003  
**LOCUS** BX282448 Soares testis\_NHT Homo sapiens cDNA clone IMAGp9981I04411  
**DEFINITION** IMAGE:1736145 - mRNA sequence.  
**ACCESSION** BX282448  
**VERSION** BX282448.1 GI:178612995  
**KEYWORDS**  
**ORGANISM** Homo sapiens (human)  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
**AUTHORS** Mammalia; Buteleostaria; Primates; Catarrhini; Hominidae; Homo.  
**1** (bases 1 to 385) Ebert,L., Heil,O., Hennig,S., Neubest,P., Partsch,B., Peters,M.,  
**2** Radefol,U., Schneider,D. and Korn,B.  
**TITLE** Human UniGeneSet - RZPD3  
**JOURNAL** Unpublished (2003)  
**COMMENT** Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGp9981I04411.  
RZPDLIB; I.M.A.G.E. CDNA Clone Collection;  
Human UniGeneSet - RZPDLIB (RZPDLIB No.972);  
http://www.rzpd.de/CloneCards/991-  
showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Fax: +49 30 32639 101  
www.rzpd.de

This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
M13r, Primer sequence: TTTCACTAGGAAACAGGTATGAC.

**FEATURES**

<b>source</b>	<b>FEATURES</b>
	Location/Qualifiers
	1..385
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGp9981I04411" ; IMAGE:1736145"
	/sex="male"
	/lab_host="DH10B"
	/clone_lib="Soares testis_NHT"
	/note="Vector: pRIT3D_Pac (Pharmacia) with a modified
	polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
	was prepared from mRNA obtained from Clontech
	Laboratories, Inc., and primed with a Not I - oligo (dT)
	Primer [5', TGTACCCATCTGAAGTGGAGGCCCGCAATTTTTTTTT 3'].
	Double-stranded cDNA was ligated to Eco RI adaptors
	(Pharmacia), digested with Not I and cloned into the Not I
	and Eco RI sites of the modified pRIT3 vector. Library
	went through one round of normalization to Cot5, and was
	constructed by Bento Soares and M. Fatima Bonaldo."

**ORIGIN**

Query Match	15.8%	Score 158, 4;	Length 385;
Best Local Similarity	99.4%	Pred. No. 1.2e-28;	
Matches	159;	Conservative 0;	Mismatches 1;
		Indels 0;	Gaps 0;

Db 471 TTACAGCGATGTTGCAATTGTAGCAGCTATAAATGCTTGAAAGAAA 530  
Db 149 TTTCAGCGTTGCAATTGTAGCAGCTATAAATGCTTGAAAGAAA 208  
Qy 531 TGTACCCAGATGATCATATTATCCTTGTGCAATTGTAGCAGCTATAAATGCTTGAAAGAAA 590  
Db 209 TGTACCCAGATGATCATATTATCCTTGTGCAATTGTAGCAGCTATAAATGCTTGAAAGAAA 268  
Qy 591 CACATCTCTCTGCTTGTGCAAGACATGACAAAGTC 630  
Db 109 CACATCTCTCTGCTTGTGCAAGACATGACAAAGTC 70  
Db 269 CACATCTCTCTGCTTGTGCAAGACATGACAAAGTC 308

RESULT 9  
**AG092191/c** AG092191 626 bp DNA linear GSS 03-NOV-2001  
**LOCUS** Pan troglodytes DNA, clone: PTB-092E12\_F, genomic survey sequence.  
**DEFINITION** AG092191  
**ACCESSION** AG092191.1 GI:16643993  
**VERSION** GSS.  
**SOURCE** Pan troglodytes (Chimpanzee)  
**ORGANISM** Pan troglodytes  
**Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;**  
**Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.**  
**REFERENCE** Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
**AUTHORS** Totoki,Y., Watanabe,H. and Sakaki,Y.  
**TITLE** BAC end sequences of Library PTB  
**JOURNAL** Unpublished  
**REFERENCE** Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
**AUTHORS** Totoki,Y., Watanabe,H. and Sakaki,Y.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (02-Aug-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama 230-0045, Japan  
(E-mail:chimbase@riken.go.jp, URL:http://hgp.gsc.riken.go.jp/),  
Tel:81-45-503-9111, Fax:81-45-503-9170  
**COMMENT** Clones are derived from the chimpanzee BAC library PTB. This BAC end  
was generated during the R&D process and may have higher chance of  
clone tracking errors.  
**PRIMERS**  
**LIBRARY** Sequencing: -21M13  
**VECTOR** Vector : PKS145  
**R.Site 1** : SacI  
**R.Site 2** : SacI  
**Location/Qualifiers**  
1..626  
1. .626  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="PTB-092E12\_F"  
/clone="PTB-092E12\_F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"  
**ORIGIN**

Query Match	15.4%	Score 155, DB 29;	Length 626;
Best Local Similarity	94.2%	Pred. No. 9.4e-28;	
Matches	161;	Conservative 0;	Indels 0;
		Gaps 0;	

Qy 215 GTTTTTCTGGATTTCTGCCAGTGTCTGCCAGCTGATGAAGGAGGTCCCC  
Db 397 GTTTTTCTGGATTTCTGCCAGTGTCTGCCAGCTGATGAAGGAGGTCCCC  
Qy 275 TGTGGTAGCTGGAGTTGGCTTGTGCTTCACTCTCTGAGG  
Db 337 TGTGGTAGCTGGAGTTGGCTTGTGCTTCACTCTCTGAGG  
338  
335 TGTGGAGCCACGGCTGATGGAGACCCAGCTCAAGAGGAAATAGGAGG 385  
Db 277 TGTGGAGCCACGGCTGATGGAGACCCAGCTCAAGAGGAAATAGGAGG 227

RESULT 10  
 H55516 Locus H55516 Chromosome 22 exon Homo sapiens linear EST 07-DEC-1995  
 DEFINITION mRNA sequence.

ACCESSION H55516.1  
 VERSION GI:1108382  
 KEYWORDS EST  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Hillier,L., Allen,M., Bowles,L., Dubugue,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schein,E., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
 AUTHORS WashU-Merck EST Project 1997  
 TITLE Unpublished (1997)  
 JOURNAL COMMENT Contact: Wilson RK Washington University School of Medicine 4440 Forest Park Parkway, Box 8501, St.. Louis, MO 63108 Tel: 314 26 1800 Fax: 314 286 1810 Email: est@watsen.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq. Primer: -28m13 rev2 ET from Amerham.

FEATURES Location/Qualifiers  
 SOURCE /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:16039131"  
 /db\_xref="taxon:9606"  
 /clone\_id="IMAGE:795702"  
 /sex="male"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker, Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5]. TGTGACCAATCTGAAGTGGAGGGCCGCCAAATTTCCTTTTTTTTTT 3'". Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN Query Match 14.6%; Score 146.4%; DB 9; Length 390;  
 Best Local Similarity 98.8%; Pred. No. 1.e-21; Indels 1; Gaps 1;  
 Matches 158; Conservative 0; Mismatches 1; Gaps 1;

Qy 471 TTACGGGATGTCATTGTAGACTTGGGCACTCATCAATTGCTGTGAAGAGAAA 530  
 Db 244 TTTCAGGATGTGCATTGTAAGCACTTGGGCACTCATCAATTGCTGTGAAGAGAAA 185

Qy 531 TGTACCCAGATGTATCATTACCTTGTGCTGAGGAGCGCGCTCTCTCAGGATTTCAGT 590  
 Db 184 TGTACCCAGATGTATCATTACCTTGTGCTGAGGAGCGCGCTCTCTCAGGATTTCAGT 126

Qy 591 CACATCTCCCTGCTTGTCCAGAACACATGACCAAGCTC 630  
 Db 125 CACATCTCCCTGCTTGTCCAGAACACATGACCAAGCTC 86

RESULT 11  
 CC325294 LOCUS CC325294  
 DEFINITION mRNA sequence.

ACCESSION AA459832  
 VERSION AA459832.1  
 KEYWORDS EST  
 SOURCE Homo sapiens (human)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE BayGenomics  
 AUTHORS BayGenomics  
 TITLE http://baygenomics.ucsf.edu/  
 JOURNAL Unpublished (2001)

COMMENT Contact: BayGenomics  
 Bay Area Functional Genomics Consortium (BayGenomics)  
 Email: info@baygenomics.ucsf.edu  
 Sequence tag generated by 5' RACE of total RNA from gene trap ES  
 cell line. ES cell lines harboring insertion mutation of target  
 gene are available upon request from BayGenomics. Annotation  
 information available from  
[http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL\\_LINE&LINE\\_ID=XML179](http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL_LINE&LINE_ID=XML179)

FEATURES Class: Gene Trap.  
 Location/Qualifiers 1. .533  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129 OLA"  
 /db\_xref="taxon:10090"  
 /sex="Male"  
 /cell\_type="Embryonic stem cell"  
 /clone\_lib="BayGenomics Gene Trap Library pGT0Lxf"  
 /note="Vector: pGT0Lxf"

ORIGIN Query Match 14.1%; Score 141.2; DB 28; Length 560;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Mismatches 23; Indels 0; Gaps 0;  
 Matches 155; Conservative

Qy 467 GAGATTACCGCATGTCATTGTAAGCACTTGAATGGCCACTATCAAATGCTGTGAGA 526  
 Db 101 GAGCACAGAACGATGTGCAATTGCAACCTGGCAACTCAAATGCTGTGAGA 160

Query Match 14.1%; Score 141.2; DB 28; Length 560;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Mismatches 23; Indels 0; Gaps 0;  
 Matches 155; Conservative

Qy 527 GAAAGTACCCAGATGTTCAAGCTTGTGCAAGGGCTGGCTTTAGGTT 586  
 Db 161 GAAGTGTACCCAGATGTTCAAGCTTGTGCAAGGGCTGGCTTTAGGTT 220

Query Match 14.1%; Score 141.2; DB 28; Length 560;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Mismatches 23; Indels 0; Gaps 0;  
 Matches 155; Conservative

Qy 587 CAGTCACATTCCTGCTTGTCAGAACATGCAAGCAAGCTCCTGAAAGATGTAAG 644  
 Db 221 TAGTCACCTCTTCCCTCTGTGTCAGAACATGCAAGCTCCTGAAAGATCAG 278

RESULT 14 CC325312 LOCUS CC325312 571 bp mRNA linear GSS 14-MAY-2003  
 XM085 BayGenomics Gene Trap Library pGT0Lxf Mus musculus cDNA, mRNA  
 DEFINITION sequence.  
 ACCESSION CC325312.1 GI:30719370  
 VERSION GSS  
 KEYWORDS SOURCE Mus musculus (house mouse)  
 ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE (bases 1 to 571)  
 AUTHORS BayGenomics.  
 TITLE http://baygenomics.ucsf.edu/  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: BayGenomics Consortium (BayGenomics)  
 Bay Area Functional Genomics Consortium (BayGenomics)  
 Email: info@baygenomics.ucsf.edu  
 Sequence tag generated by 5' RACE of total RNA from gene trap ES  
 cell line. ES cell lines harboring insertion mutation of target  
 gene are available upon request from BayGenomics. Annotation  
 information available from  
[http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL\\_LINE&LINE\\_ID=XM085](http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL_LINE&LINE_ID=XM085).

FEATURES Class: Gene Trap.  
 Location/Qualifiers 1. .571  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129 OLA"  
 /db\_xref="taxon:10090"  
 /sex="Male"  
 /clone\_lib="BayGenomics Gene Trap Library pGT0Lxf"  
 /note="Vector: pGT0Lxf"

ORIGIN Query Match 14.1%; Score 141.2; DB 28; Length 571;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Mismatches 23; Indels 0; Gaps 0;

Qy 467 GAGATTACCGCATGTCATTGTAAGCACTTGAATGGCCACTATCAAATGCTGTGAGA 526  
 Db 112 GAGCACAGAACGATGTGCAATTGCAACCTGGCAACTCAAATGCTGTGAGA 171

Query Match 14.1%; Score 141.2; DB 28; Length 571;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Mismatches 23; Indels 0; Gaps 0;

Qy 527 GAAAGTACCCAGATGTTCAAGCTTGTGCAAGGGCTGGCTTTAGGTT 586  
 Db 172 GAAGTGTACCCAGATGTTCAAGCTTGTGCAAGGGCTGGCTTTAGGTT 231

Query Match 14.1%; Score 141.2; DB 28; Length 571;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Mismatches 23; Indels 0; Gaps 0;

Qy 587 CAGTCACATTCCTGCTTGTCAGAACATGCAAGCAAGCTCCTGAAAGATGTAAG 644  
 /cell\_type="Embryonic stem cell"

Db	232	TAGTCACTTCCTCTTCCTGTCAGAACACATCGACCAAGCTCTGAAGATCAAAG	289		Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
RESULT	15			COMMENT	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
AK054270	AK054270	Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone: b33008K23 product: hypothetical HMG-I and HMG-Y DNA-binding domain (A+T-hook) PHD-finger/DHHC-type Zn-finger/RING finger containing protein, full insert sequence.			Please visit our web site for further details. URL: http://genome-gsc.riken.go.jp/
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE				FEATURES	
AUTHORS				SOURCE	
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasakai, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Harada, A., Isono, N., Ishii, Y., Nakamura, S., Hazama, M., Nishime, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwasa, M., Ohara, E., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, X., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system -3.64-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the RIKEN Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
REFERENCE	5				
AUTHORS	The RIKEN Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.				
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
JOURNAL	Nature 420, 563-573 (2002)				
REFERENCE	6 (bases 1 to 2647)				
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Adachi, S., Furuno, M., Hanagata, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirono, T., Hori, F., Imotani, K., Ishii, Y., Iton, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Onsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Takeda, Y., Tanaka, T., Tomaru, A., Tagawa, A., Takahashi, F., Toku, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (16 JULY 2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)				

Search completed: July 9, 2004, 10:07:38  
Job time : 4200 secs

This Page Blank (uspto)

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 16:57:41 ; Search time 55 Seconds  
 (without alignments)

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: geneseqD1980s:\*
- 2: geneseqD1990s:\*
- 3: geneseqD000s:\*
- 4: geneseqD2001s:\*
- 5: geneseqD2002s:\*
- 6: geneseqD2003abs:\*
- 7: geneseqD2003bs:\*
- 8: geneseqD2004s:\*

Total number of hits satisfying chosen parameters: 1586107

26 62 29.1 384 4 AAM40180 Human pol  
 27 62 29.1 384 4 Aab94589 Human pro  
 28 62 29.1 384 4 Aab93955 Human pro  
 29 62 29.1 384 4 AAE00924 Human sph  
 30 62 29.1 384 4 AAE07882 Human sph  
 31 62 29.1 384 5 ABG31586 Human sph  
 32 62 29.1 384 7 Ade38365 Human pro  
 33 60 28.2 552 6 Abu47220 Protein e  
 34 58 27.2 173 4 ABB71060 Drosophil  
 35 58 27.2 384 5 ABB08089 Human sph  
 36 57 27.0 340 6 ABU44926 Protein e  
 37 57 26.8 398 2 AAR78673 CD4 domain  
 38 57 26.8 398 2 AAR89450 CD4 DI-D4  
 39 57 26.8 399 1 AAP93010 Genetic C  
 40 57 26.8 416 3 AAB19509 CD4-IGM f  
 41 57 26.8 432 2 AAR74222 Epitope O  
 42 57 26.8 432 2 AAY30514 Predicted  
 43 57 26.8 432 6 ADA25188 CD4 epitope  
 44 57 26.8 433 2 AAW41376 Human CD4  
 45 57 26.8 433 3 AAY54500 Amino aci

## ALIGNMENTS

RESULT 1  
 AAW02152 ID AAW02152 standard; Protein; 43 AA.  
 XX AC  
 XX DT 04-DEC-1996 (First entry)  
 XX DE BAGE tumour rejection antigen precursor.  
 XX KW tumour rejection antigen precursor; TRAP; MHC;  
 KW major histocompatibility complex; HLA-Cw\*1601; melanoma; metastasis;  
 KW diagnosis; therapy; vaccine.  
 XX Homo sapiens.  
 XX CS  
 XX Key  
 FT Location/Qualifiers  
 FT 2..10  
 FT /label= Tumour\_rejection\_antigen

WO9625511-A1.

XX PD 22-AUG-1996.  
 XX PF 07-FEB-1996; 96WO-US001608.  
 XX PR 16-FEB-1995; 95US-00389360.  
 XX (LUDWIG INST CANCER RES.  
 XX PA  
 XX PI Boel P, Wildmann C, Boonfalleur T, Van Der Bruggen P, Coulie P;  
 PI Renaud J;  
 XX XX WPI: 1996-393411/39.  
 DR DR - N-PSDB; AAT36332.

Tumour rejection antigen precursor (TRAP) and gene - useful to develop  
 prods. for diagnosis and treatment of disorders characterised by TRAP,  
 partic. melanoma.

Example 4; Page 11; 44DP; English.

The BAGE tumour rejection antigen precursor (TRAP) (AAW02152) is  
 processed to a tumour rejection antigen (TRA) (AAW02153) presented by the  
 MHC molecule HLA-Cw1601. The amino acid sequence of the BAGE TRAP was  
 deduced from a cDNA clone (AAT36332) derived from melanoma cell line MZ2-  
 MEL 43. BAGE expression was not observed in healthy adult or foetal  
 tissue, but was detected in 22% of melanoma lines examined, being partic.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	213	100.0	43	2	AAW02152	BAGE tumo	Aaw02152 Human BAG
2	213	100.0	43	5	AAU84811	Human BAG	Aau84811 Human BAG
3	213	100.0	43	5	ABB78347	Amino aci	Abb78347 Amino aci
4	213	100.0	43	7	ADD25520	Birdwing d	Add25520 Birdwing d
5	156	73.2	30	5	AAU85010	Human BAG	Aau85010 Human BAG
6	156	73.2	30	5	AAU85130	Human mel	Aau85130 Human mel
7	126	59.2	30	5	AAU85009	Human BAG	Aau85009 Human BAG
8	98	46.0	22	2	AAR67808	BAGE tumo	Aar67808 BAGE tumo
9	98	46.0	22	2	AAY10634	Peptide a	Aay10634 Peptide a
10	98	46.0	22	5	ABG83319	MHC Class	Abg83319 MHC Class
11	87	40.8	17	5	AAU85011	Human BAG	Aau85011 Human BAG
12	70	32.9	16	5	AAR67809	BAGE tumo	Aar67809 BAGE tumo
13	70	32.9	16	2	AAV10635	Peptide a	Aav10635 Peptide a
14	70	32.9	16	5	ABG83316	MHC Class	Abg83316 MHC Class
15	63	29.6	384	4	AAB48007	Human sph	Aab48007 Human sph
16	63	29.6	384	6	ABP71054	Human sph	Abp71054 Human sph
17	62	29.1	293	2	AAW84613	Secreted	Aaw84613 Secreted
18	62	29.1	293	4	ABBS0380	Human sec	Abbs0380 Human sec
19	62	29.1	293	6	ABO44637	Novel hum	Abo44637 Novel hum
20	62	29.1	293	7	ABO26117	Human pro	Abo26117 Human pro
21	62	29.1	305	4	AAM41966	Human pol	Aam41966 Human pol
22	62	29.1	333	5	ABBS0209	Human pol	Abbs0209 Human pol
23	62	29.1	368	6	ABR2391	Human sph	Abr2391 Human sph
24	62	29.1	384	3	ABJ18659	A human r	Abj18659 A human r
25	62	29.1	384	3	AYA96057	Human sph	Aya96057 Human sph

CC prevalent in metastatic lesions. The BAGE antigen can be expressed in transformed or transfected host cells. It is useful for identifying those cells which present HLA-Cw1601 and may be incorporated into vaccines useful in treating disorders characterised by expression of the TRAP molecule

XX Sequence 43 AA;

Query Match	100.0%	Score 213;	DB 2;	Length 43;
Best Local Similarity	100.0%	Pred. No. 1.	8e-23;	
Matches	43;	Conservative	0;	Gaps 0;
Qy	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCFF 43			
Db	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCFF 43			

RESULT 2  
AAU84811

ID AAU84811 standard; protein; 43 AA.

XX AC AAU84811;

XX DT 08-MAY-2002 (first entry)

XX DE Human BAGE consensus sequence.

XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus; viral infection; human immunodeficiency virus; melanoma; bacterial infection; Salmonella; Legionella; parasitic infection; Trypanosoma; Toxoplasma; Giardia.  
XX OS Homo sapiens.

XX PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU000622.

XX PR 26-MAY-2000; 2000AU\_000007761.

XX PA (AUSU ) UNIV AUSTRALIAN NAT.

XX PI Thomson SA, Ramshaw IA;

XX DR WPI; 2002-147575/19.

XX PT New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.

XX PS Example 3; Fig 27; 364pp; English.

XX The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as potentiating agents. Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis virus, Epstein-Barr virus and respiratory syncytial virus), bacterial (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,

CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic (e.g., infections caused by Plasmodium, Schistosoma, Leishmania, Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is a consensus sequence for a parent protein used to design a Savine of the invention

XX SQ Sequence 43 AA;

Query Match	100.0%;	Score 213;	DB 5;	Length 43;
Best Local Similarity	100.0%;	Pred. No. 1.	8e-23;	
Matches	43;	Conservative	0;	Gaps 0;
Qy	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCFF 43			
Db	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCFF 43			

RESULT 3  
ABB73347

ID ABB73347 standard; protein; 43 AA.

XX AC ABB73347;

XX DT 16-DEC-2002 (first entry)

XX DE Amino acid sequence of human BAGE.

XX KW Human; IFIT-2; chronic myelogenous leukemia; IAGE-1; BAGE; DDB1; ETS1; PIASy; PIASx-alpha; PIASx-beta; DAPK3.  
XX OS Homo sapiens.

XX PN WO200270747-A1.

XX PD 12-SSP-2002.

XX PR 01-MAR-2002; 2002WO-JP001901.

XX PR 01-MAR-2001; 2001JP-00056438.

XX PA (FUJII ) FUJISAWA PHARM CO LTD.

XX PI Mano H;

XX WPI; 2002-682911/73.

DR N-PSB; ABV72225.

XX PT Measuring the expression profile of genes in a cell or tissue sample for diagnosis of chronic myelogenous leukemia and identification of agents for its treatment.

XX PS Example 3; Page 39-40; 97pp; Japanese.

XX CC The present sequence represents human BAGE. The expression level of the gene is used in the method of the invention. The specification describes a method of examining chronic myelogenous leukemia. The method comprises measuring the expression level of a gene selected from IFIT-2, IAGE-1, BAGE, DDB1, ETS2, PIASy, PIASx-alpha, PIASx-beta and DAPK3, or determining the expression profile of a group of genes including one or more of these genes, in a cell or tissue sample from a chronic myelogenous leukemia patient. The method is used for the diagnosis, treatment and prevention of chronic myelogenous leukemia

XX SQ Sequence 43 AA;

Query Match	100.0%;	Score 213;	DB 5;	Length 43;
Best Local Similarity	100.0%;	Pred. No. 1.	8e-23;	
Matches	43;	Conservative	0;	Gaps 0;
Qy	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCFF 43			
Db	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCFF 43			

**RESULT 4**  
**ADD2520 standard; protein; 43 AA.**

XX ADD2520;  
 ID XX  
 AC XX  
 DT 15-JAN-2004 (first entry)  
 DE Binding domain-immunoglobulin fusion protein-associated protein #37.  
 KW Binding domain; immunoglobulin; fusion protein; cytosatic;  
 KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
 KW neuroprotective; hinge region; immunoglobulin heavy chain;  
 KW CH2 constant region; CH3 constant region; IgG1;  
 KW malignant dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
 KW B-cell disorder; melanoma; carcinoma; sarcoma;  
 KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
 KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
 XX Unidentified.  
 OS US2003118592-A1.  
 PN XX  
 PD 26-JUN-2003.  
 XX PR 25-JUL-2002; 2002US-00207655.  
 XX PR 17-JAN-2001; 2001US-0367358P.  
 PR 03-JUN-2002; 2002US-00053350.  
 PR 03-JUN-2002; 2002US-038561P.  
 XX PA (GENE-) GENECRAFT INC.  
 PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
 XX DR 2003-801317/75.  
 XX PT New binding domain-immunoglobulin fusion protein, useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
 PS Disclosure: SEQ ID NO 81; 157pp; English.  
 XX CC The invention relates to a binding domain-immunoglobulin fusion protein comprising a binding domain polypeptide that is fused to an immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain CH2 constant region polypeptide that is fused to the hinge region polypeptide, and an immunoglobulin heavy chain CH3 constant region polypeptide that is fused to the CH2 constant region polypeptide. The hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin hinge region polypeptide, a mutated human IgG1 immunoglobulin hinge region polypeptide, derived from (a) having 3 or more cysteine residues; where the mutated human IgG1 immunoglobulin hinge region polypeptide contains 2 cysteine residues, where the first cysteine is not mutated; a mutated human IgG1 immunoglobulin hinge region polypeptide, derived from (a) having 3 or more cysteine residues, where the mutated human IgG1 immunoglobulin hinge region polypeptide contains no more than one cysteine residue; and a mutated human IgG1 immunoglobulin hinge region polypeptide comprising antibody dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The binding domain polypeptide is capable of specifically binding to an antigen. Also included are an isolated polynucleotide encoding the binding domain-immunoglobulin fusion protein, a recombinant expression construct comprising the polynucleotide operably linked to a promoter, a host cell transformed or transfected with a recombinant expression construct, producing the binding domain-immunoglobulin fusion protein, a pharmaceutical composition comprising the binding domain-immunoglobulin fusion protein or polynucleotide and a carrier, and treating a subject having or suspected of having a malignant condition or a B-cell disorder.

CC The binding domain-immunoglobulin fusion protein is useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoïd arthritis, myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple sclerosis or autoimmune disease. The present sequence is a binding domain -immunoglobulin fusion protein-associated protein sequence. Note: The sequence data for this patent formed part of the printed specification and is also available in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not identified the sequences in the printed specification by their SEQ ID number therefore none of the sequences can be explicitly identified.

XX SQ Sequence 43 AA;

Query	Match	Score 100.0%	Length 43;
Best Local Similarity	100.0%	Pred. No. 1.8e-23;	
Matches	43;	Conservative 0;	Mismatches 0;
Indels	0;	Gaps 0;	

XX Qy 1 MAARAVFLAQLAQQLQARLMKEESPVSVWLEPEDTALCFIF 43  
 Db 1 MAARAVFLAQLAQQLQARLMKEESPVSVWLEPEDTALCFIF 43

RESULT 5  
 AAU85010  
 ID AAU85010 standard; peptide; 30 AA.  
 XX XX  
 AC AAU85010;  
 XX XX  
 DT 08-MAY-2002 (first entry)  
 DE Human BAGE segment 2.  
 XX XX  
 Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; Toxoplasma; parasitic infection;  
 KW Trypanosoma; Giardia.  
 XX OS Homo sapiens.  
 XX XX  
 FN WO200190197-A1.  
 XX XX  
 PP 25-MAY-2001; 2001WO-AU000622.  
 XX XX  
 PR 26-MAY-2000; 2000AU-00007761.  
 XX PA (AUSU ) UNIV AUSTRALIAN NAT.  
 XX PI Thomson SA, Ramshaw IA;  
 XX DR WPI; 2002-14755/19.  
 XX XX  
 PT New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.  
 PS Example 3; Fig 27; 364pp; English.  
 XX XX  
 PT New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.  
 PS Example 3; Fig 27; 364pp; English.  
 XX XX  
 PT The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relative position to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for modulating immune responses preferentially directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head,

CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as potentiating agents. Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HRV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis (e.g., infections caused by *Neisseria*, *Meningococcal*, *Haemophilus*, *Salmonella*, *Streptococcal*, *Legionella* and *Mycobacterium* or parasitic (e.g., infections caused by *Plasmodium*, *Schistosoma*, *Leishmania*, *Trypanosoma*, *Toxoplasma* and *Giardia*) infections. The present sequence is a peptide derived from a parent protein used to construct a Savine of the invention.

XX Sequence 30 AA;

Query Match 73.2%; Score 156; DB 5; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2.5e-15; Indels 0; Gaps 0;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LLQARLMKEESPVYSWRLSPEDGTALCFIF 43  
Db 1 LLQARLMKEESPVYSWRLSPEDGTALCFIF 30

## RESULT 6

AAU85130

ID AAU85130 standard; protein; 3541 AA.

XX

AC AAU85130;

XX

DT 08-MAY-2002 (first entry)

XX

DE Human melanoma specific savine.

XX

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

KW viral infection; human immunodeficiency virus; melanoma;

KW bacterial infection; *Salmonella*; *Legionella*; parasitic infection;KW *Trypanosoma*; *Toxoplasma*; *Giardia*.

XX OS Homo sapiens.

OS Synthetic.

XX PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PP 25-MAY-2001; 2001WO-AU000622.

XX PR 26-MAY-2000; 2000AU-00007761.

XX PA (AUSU ) UNIV AUSTRALIAN NAT.

PA XX PI Thomson SA, Ramshaw IA;

PI XX DR WPI; 2002-147575/19.

DR N-PSDB; ABK36839.

XX New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.

XX PS Example 3; Fig 27; 364pp; English.

XX

The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for

CC modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus) as potentiating agents. Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis (e.g., infections caused by *Neisseria*, *Meningococcal*, *Haemophilus*, *Salmonella*, *Streptococcal*, *Legionella* and *Mycobacterium* or parasitic (e.g., infections caused by *Neisseria*, *Meningococcal*, *Haemophilus*, *Salmonella*, *Streptococcal*, *Legionella* and *Mycobacterium*, *Trypanosoma*, *Schistosoma*, *Leishmania*, *Trypanosoma*, *Toxoplasma* and *Giardia*) infections. The present sequence is a Savine protein of the invention

XX

SQ Sequence 30 AA;

Query Match 73.2%; Score 156; DB 5; Length 3541;  
Best Local Similarity 100.0%; Prod. No. 6.1e-13;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LLQARLMKEESPVYSWRLSPEDGTALCFIF 43  
Db 2777 LLQARLMKEESPVYSWRLSPEDGTALCFIF 2806

## RESULT 7

AAU85009

ID AAU85009 standard; peptide; 30 AA.

XX AAU85009;

AC AAU85009;

XX DT 08-MAY-2002 (first entry)

XX DE Human BAGE segment 1.

XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

KW viral infection; human immunodeficiency virus; melanoma;

KW bacterial infection; *Salmonella*; *Legionella*; parasitic infection;KW *Trypanosoma*; *Toxoplasma*; *Giardia*.

XX OS Homo sapiens.

OS PN WO200190197-A1.

XX PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU000622.

XX PR 26-MAY-2000; 2000AU-00007761.

XX PA (AUSU ) UNIV AUSTRALIAN NAT.

PA XX PI Thomson SA, Ramshaw IA;

PI XX DR WPI; 2002-147575/19.

DR N-PSDB; ABK36839.

XX New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.

XX PS Example 3; Fig 27; 364pp; English.

XX

The invention relates to new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for

CC The invention relates to new synthetic polypeptide (I) comprising CC several different segments of at least one parent polypeptide linked CC together in a different relationship relative to their linkage in the CC parent polypeptide to impede, abrogate or otherwise alter at least one CC function associated with the parent polypeptide and for inducing an CC immune response against a pathogen or cancer. Also included are a CC synthetic polynucleotide encoding and a computer system for designing the CC synthetic polypeptides. The synthetic polypeptides and polynucleotides CC are referred to as a Savine. The synthetic polypeptide is useful for CC

CC modulating immune responses preferably directed against a pathogen or a cancer. (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis virus, Epstein-Barr virus and respiratory syncytial virus), bacterial (e.g., infections caused by Neisseria, Meningococcal, Haemophilus, Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is a peptide derived from a parent protein used to construct a savine of the invention

XX Sequence 30 AA;

Query Match 59.2%; Score 126; DB 5; Length 30;  
Best Local Similarity 100.0%; Pred. No. 6.1e-11;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLAISQQLQARLMKE 28  
DB 3 MAARAVFLAISQQLQARLMKEESPVVS 30

RESULT 8

ID AAR67808 standard; peptide; 22 AA.  
XX

AC AAR67808;

XX

DT 25-MAR-2003 (revised)

DT 22-AUG-1995 (first entry)

XX DE BAGE tumor rejection antigen peptide.

XX KW BAGE: tumor rejection antigen precursor; diagnosis; HLA;

KW human leukocyte antigen MHC; major histocompatibility complex; TRAP; cancer; melanoma.

XX OS Synthetic.

XX PN WO9500159-A1.

XX PD 05-JAN-1995.

XX PF 10-JUN-1994; 94WO-US006534.

XX PR 17-JUN-1993; 93US-00079110.

XX PR 15-FEB-1994; 94US-00196630.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Van Der Bruggen P, Boon-Falleur T, Coulie P, Renaud J;

XX DR WPI: 1995-051741/07.

Nucleic acid coding for a tumour rejection antigen precursor - used to PT develop prods. for the diagnosis and therapy of cancers, partic. PR melanomas.

XX PS Claim 21; Page 19; 33pp; English.

This sequence encodes the tumor rejection antigen peptide BAGE. The peptide may be used in the diagnosis and therapy of cancers, e.g. CC melanomas. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 22 AA;

Query Match 46.0%; Score 98; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

QY 1 MAARAVFLAISQQLQARLMKE 22  
DB 1 MAARAVFLAISQQLQARLMKE 22

RESULT 9

AY10634  
ID AY10634 standard; peptide; 22 AA.

XX

AC AAY10634;

XX

DT 12-MAY-1999 (first entry)

XX

DB Peptide antigen SEQ ID NO: 564.

XX

CY Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system; immunisation; tumour; infectious disease; immunotherapy; cancer; malignant melanoma; viral disease; hepatitis; AIDS.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9002183-A2.

XX

PD 21-JAN-1999.

XX

PF 10-JUL-1998; 98WO-US014289.

XX

PR 10-JUL-1997; 97CA-02209815.

XX

PR 10-DEC-1997; 97US-00988320.

XX

PA (CTL-) CTL IMMUNOTHERAPIES CORP.

XX

P1 Kuendig TM, Simard JJL;

XX

DR WPI: 1999-120514/10.

XX

Inducing a cytotoxic T lymphocyte response - by maintaining a level of PT antigen in the lymphatic system of a mammal so as to provide a sustained PT CTL response, used to treat, e.g. AIDS.

XX

PS Disclosure; Page 52; 199pp; English.

XX

DR WPI: 1999-120514/10.

XX

The present invention describes a method of inducing and/or sustaining an CC immunological cytotoxic T lymphocyte (CTL) response in a mammal. The CC method comprises: (a) delivering an antigen to the mammal at a level to CC induce an immunological CTL response in the mammal; and (b) maintaining CC the level of the antigen in the mammal's lymphatic system to maintain the CC immunologic CTL response. The method can be used for the delivery of e.g. CC a differentiation antigen, a tumour-specific multilinage antigen, an CC embryonic antigen, an oncogene antigen, a mutated tumour-suppressor gene CC antigen, or a viral antigen. They can be used for the treatment of CC disease such as cancer, e.g. malignant melanoma or infectious disease, CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery CC to the lymphatic system provides for potent CTL stimulation that takes CC place in the milieu of the lymphoid organ, and its subsains stimulation CC that is necessary to keep CTL active, cytotoxic and recirculating through CC the body. AAY10071 to AAY10639 represent examples of peptide antigens CC given in the present invention

XX

Sequence 22 AA;

QY 1 MAARAVFLAISQQLQARLMKE 22

DB 1 MAARAVFLAISQQLQARLMKE 22

XX

Query Match 46.0%; Score 98; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

ABG80319 standard; peptide; 22 AA.  
 XX DE Human BAGE segment 3.  
 ID ABG80319; XX  
 AC ABG80319; XX  
 XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 XX KW viral infection; human immunodeficienty virus; melanoma;  
 XX bacterial infection; Salmonella; Legionella; parasitic infection;  
 DT 29-AUG-2003 (revised) KW TRYPANOSOMA; Toxoplasma; Giardia.  
 DT 15-NOV-2002 (first entry) XX  
 DB MHC class I molecule, viral epitope #567. XX  
 KW Major histocompatibility complex; MHC class I molecule; virus; XX  
 KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system; XX  
 KW antigen; immunogenic; malignant tumour; carcinoma; melanoma; leukaemia; XX  
 KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis; XX  
 KW acquired immune deficiency syndrome; AIDS. XX  
 OS Viruses. OS Homo sapiens.  
 XX PA (AUSU ) UNIV AUSTRALIAN NAT.  
 PN WO2002622368-A2. XX  
 PD 15-AUG-2002. PI Thomson SA, Ramshaw IA;  
 XX PP 22-JAN-2002; 2002WO-US002033. XX  
 PR 02-FEB-2001; 2001US-00776232. DR WPI; 2002-147575/19.  
 XX PA (CTL1-) CTL IMMUNOTHERAPIES CORP. DR N-PSDB; ABK36831.  
 PR 02-FEB-2001; 2001US-00776232. XX  
 PT New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the parent polypeptide, for inducing immune response against a pathogen or cancer. PT  
 PT Example 3; Fig 27; 364pp; English. PT  
 PS XX  
 CC The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for modulating immune responses preferably directed against a pathogen or a cancer. (e.g., cancers of the lung, breast, ovary, cervix, colon, head, neck, pancreas, prostate, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as potentiating agents. Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis virus, Epstein-Barr virus and respiratory syncytial virus), bacterial (e.g., infections caused by Neisseria, Meningococcal, Haemophilus, Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic (e.g., infections caused by Plasmodium, Schistosoma, Leishmania, Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is a peptide derived from a parent protein used to construct a savine of the invention XX  
 CC Sequence 17 Å;  
 SQ XX  
 CC Query Match 40.1%; Score 87; DB 5; Length 17;  
 CC Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 Qy 29 WRLPEDGTALCFIF 43  
 Db 1 WRLPEDGTALCFIF 15

RESULT 11  
 ABU85011  
 ID AAU85011 standard; peptide; 17 AA.  
 XX AC AAU85011;  
 XX DT 08-MAY-2002 (first entry)

RESULT 12  
 AAR67809  
 ID AAR67809 standard; peptide; 16 AA.  
 XX AC AAR67809;  
 XX AC AAR67809;  
 XX DT

DT 25-NAR-2003 (revised)  
 DT 22-AUG-1995 (first entry)  
 XX BAGE tumor rejection antigen peptide.  
 DE XX  
 KW BAGE; tumor rejection antigen precursor; diagnosis; HLA;  
 human leukocyte antigen MHC; major histocompatibility complex; TRAP;  
 cancer; melanoma.  
 XX Synthetic.  
 OS XX  
 PN WO9500159-A1.  
 XX  
 PD 05-JAN-1995.  
 XX PR 10-JUN-1994; 94WO-US006534.  
 XX PR 17-JUN-1993; 93US-00079110.  
 XX PR 15-FEB-1994; 94US-00196630.  
 XX PA (LUDWIG) LUDWIG INST CANCER RES.  
 XX PI Van Der Bruggen P, Boon-Falleur T, Coulie P, Renaud J;  
 XX WPI: 1995-051741/07.  
 XX Nucleic acid coding for a tumour rejection antigen precursor - used to  
 PT develop prods. for the diagnosis and therapy of cancers, partic.  
 PT melanomas.  
 XX  
 PS Claim 21; Page 20; 33pp; English.  
 XX  
 CC This sequence encodes the tumor rejection antigen Peptide BAGE. The  
 CC peptide may be used in the diagnosis and therapy of cancers, e.g.  
 CC melanomas. (Updated on 25-MAR-2003 to correct PN field.)  
 XX SQ Sequence 16 AA;  
 XX  
 Query Match 32.9%; Score 70; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.0045;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1 MAARAVFLAQLQ 16  
 AC ||||| ||||| |||||  
 DB 1 MAARAVFLAQLQ 16  
 XX  
 RESULT 13  
 AAY10635 ID AAY10635 standard; peptide; 16 AA.  
 XX  
 AC AAY10635;  
 DT 12-MAY-1999 (first entry)  
 DE Peptide antigen SEQ ID NO:565.  
 XX  
 KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;  
 KW immunisation; tumour; infectious disease; immunotherapy; cancer;  
 KW malignant melanoma; viral disease; hepatitis; AIDS.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX PN WO9902183-A2.  
 XX PR 21-JAN-1999.  
 XX PF 10-JUL-1998; 98WO-US014289.  
 XX PR 10-JUL-1997; 97CA-02209815.  
 XX PR 10-DEC-1997; 97US-0098320.  
 XX  
 PA (CTLI-) CTL IMMUNOTHERAPIES CORP.  
 XX PI Kundiig TM, Simard JJL;  
 DR WPI: 1999-120514/10.  
 XX  
 Inducing a cytotoxic T lymphocyte response - by maintaining a level of  
 PT antigen in the lymphatic system of a mammal so as to provide a sustained  
 PT CTL response, used to treat, e.g. AIDS.  
 XX Disclosure; Page 52; 199pp; English.  
 XX  
 The present invention describes a method of inducing and/or sustaining an  
 CC immunological cytotoxic T lymphocyte (CTL) response in a mammal. The  
 CC method comprises: (a) delivering an antigen to the mammal at a level to  
 CC induce an immunological CTL response in the mammal's lymphatic system to maintain the  
 CC level of the antigen in the mammal's lymphatic system to maintain the  
 CC immunologic CTL response. The method can be used for the delivery of e.g.  
 CC a differentiation antigen, a tumour-specific multipeptide antigen, an  
 CC embryonic antigen, an oncogene antigen, a mutated tumour-suppressor gene  
 CC antigen, or a viral antigen. They can be used for the treatment of  
 CC disease such as cancer, e.g. malignant melanoma or infectious disease,  
 CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery  
 CC to the lymphatic system provides for potent CTL stimulation that takes  
 CC place in the milieu of the lymphoid organ, and it sustains stimulation  
 CC that is necessary to keep CTL active, cytotoxic and recirculating through  
 CC the body. AAY10671 to AAY10639 represent examples of peptide antigens  
 CC given in the present invention  
 XX SQ Sequence 16 AA;  
 XX  
 Query Match 32.9%; Score 70; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.0045;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1 MAARAVFLAQLQ 16  
 AC ||||| ||||| |||||  
 DB 1 MAARAVFLAQLQ 16  
 XX  
 RESULT 14  
 ABG80316 ID ABG80316 standard; peptide; 16 AA.  
 XX AC ABG80316;  
 XX PR 29-AUG-2003 (revised)  
 DT 15-NOV-2002 (first entry)  
 DE MHC class I molecule, viral epitope #564.  
 XX  
 KW Major histocompatibility complex; MHC; MHC class I molecule; virus;  
 KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system;  
 KW antigen; immunogenic; malignant tumour; carcinoma; melanoma; leukaemia;  
 KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis;  
 KW acquired immune deficiency syndrome; AIDS.  
 XX Viruses.  
 XX PN WO200262368-A2.  
 XX PD 15-AUG-2002.  
 XX PF 22-JAN-2002; 2002WO-US002033.  
 XX PR 02-FEB-2001; 2001US-00776232.  
 XX  
 PA (CTLI-) CTL IMMUNOTHERAPIES CORP.  
 XX PI Kundiig TM, Simard JJL;  
 DR WPI: 2002-657506/70.

PT Inducing or sustaining immunological cytotoxic T lymphocyte response in a mammal, useful for treating a malignant tumor or infectious disease, by directly administering an antigen to the lymphatic system of the mammal.

XX Disclosure: Page 46; 73PP; English.

XX The invention relates to a method of inducing and/or sustaining an immunological cytotoxic T lymphocyte (CTL) response in a mammal comprising administering directly to the lymphatic system of the mammal: (a) an antigen in the form of a polypeptide; (b) a vector comprising a nucleic acid encoding the antigen; or (c) a non-peptide antigen. The method is useful for inducing and/or sustaining CTL response in a mammal. This is particularly useful for creating a mammal having a malignant tumour (e.g. carcinoma, melanoma, leukaemia or lymphoma) or infectious disease (e.g. hepatitis, acquired immune deficiency syndrome (AIDS), malaria, measles or tuberculosis), or in an animal having a predisposition to these diseases. The mammal may be dogs, cats, mice, cattle, sheep, pigs, goats, rabbits, or preferably humans. ABG79733-ABG0319 represent viral epitopes on major histocompatibility complex (MHC) class I molecules used in the method of the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 16 AA;

Query Match 32.9%; Score 70; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0045; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAISAQIQLQ 16

Db 1 MAARAVFLAISAQIQLQ 16

#### RESULT 15

ID AAB48007 standard; protein; 384 AA.  
XX AAB48007;

AC XX

DT 19-MAR-2001 (First entry)

DE Human Sphingosine kinase protein sequence.

XX Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic; antiinflammatory; neuroprotective; antibacterial; immunosuppressive; human.

OS Homo sapiens.

XX PN WO200070028-A1.

XX

PD 23-NOV-2000.

PF 12-MAY-2000; 2000WO-AU000457.

XX PR 13-MAY-1999; 99AU-00000339.

PR 08-JUL-1999; 99AU-00001504.

XX PA (JOHJ ) JOHNSON & JOHNSON RES PTY LTD.

XX Pitson SM, Wattenberg BW, Xia P, D'andrea RJ, Gamble JR;  
PI Vadas MA;

XX

DR WPI; 2001-016227/02.

DR N-PSDB; AAC84161.

XX Novel sphingosine kinase protein and nucleic acid molecules for diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma, atherosclerosis, inflammation, meningitis, multiple sclerosis and septic shock.

XX Claim 9; Fig 7a; 100pp; English.

XX This represents a human sphingosine kinase (SK) protein. The human SK protein, encoding nucleic acids and modulators are useful for modulating expression, functional activity or cellular functional activity of sphingosine kinase in a subject and also for treating a mammal by modulating the activity of SK. Diseases treated by regulating SK cellular activity include rheumatoid arthritis, asthma, atherosclerosis, inflammation, meningitis, multiple sclerosis and septic shock.

XX Query Match 29.6%; Score 63; DB 4; Length 384;  
Best Local Similarity 37.2%; Pred. No. 1.8;  
Matches 16; Conservatve 9; Mismatches 12; Indels 6; Gaps 1;

Qy 3 ARAVFALSAQLLQRLMKES-----PVVSWLEPEDGTAL 39  
Db 295 SRAMLLRLFLAMEKGRHMEYECPILVYYVPAFRLPEPDGKGM 337

Search completed: July 7, 2004, 17:07:44  
Job time : 57 secs

SQ Sequence 384 AA;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw mode1  
Run on: July 7, 2004, 17:06:43 ; Search time 23 Seconds  
(without alignments)  
96.518 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213

Sequence: 1 MAARAVFLAASQQLQARLM.....SPVVSWRLEPEDGFTALCFIF 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters:

389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgcn2\_6/ptodata/2/iaa/5A\_COMBO.PEP:\*

2: /cgcn2\_6/ptodata/2/iaa/5B\_COMBO.PEP:\*

3: /cgcn2\_6/ptodata/2/iaa/6A\_COMBO.PEP:\*

4: /cgcn2\_6/ptodata/2/iaa/6B\_COMBO.PEP:\*

5: /cgcn2\_6/ptodata/2/iaa/PCTUS\_COMBO.PEP:\*

6: /cgcn2\_6/ptodata/2/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	43	2 US-08-389-360-2	Sequence 2, Appli
2	213	100.0	43	3 US-09-038-328-2	Sequence 2, Appli
3	213	100.0	43	4 US-09-435-24-2	Sequence 2, Appli
4	213	100.0	43	4 US-09-382-197-2	Sequence 2, Appli
5	98	46.0	14	1 US-08-196-630A-8	Sequence 9, Appli
6	70	32.9	16	1 US-08-196-630A-9	Sequence 9, Appli
7	62	29.1	54	4 US-09-621-576-5718	Sequence 328, AD
8	62	29.1	293	4 US-09-025-58-328	Sequence 328, AD
9	62	29.1	384	4 US-09-970-516-2	Sequence 2, Appli
10	57	26.8	398	2 US-08-284-391B-29	Sequence 29, Appli
11	57	26.8	398	3 US-09-218-550-29	Sequence 29, Appli
12	57	26.8	433	2 US-08-867-149-1	Sequence 1, Appli
13	57	26.8	433	2 US-08-808-374-1	Sequence 1, Appli
14	57	26.8	433	3 US-09-100-109A-1	Sequence 1, Appli
15	57	26.8	458	3 US-08-466-368-4	Sequence 4, Appli
16	57	26.8	462	2 US-08-417-495-5	Sequence 5, Appli
17	57	26.8	462	2 US-08-284-391B-5	Sequence 5, Appli
18	57	26.8	462	3 US-09-218-950-5	Sequence 5, Appli
19	57	26.8	462	5 PCT-US92-01785-6	Sequence 5, Appli
20	57	26.8	462	5 PCT-US95-00454-6	Sequence 4, Appli
21	57	26.8	532	2 US-08-417-495-6	Sequence 6, Appli
22	57	26.8	532	2 US-08-417-495-6	Sequence 6, Appli
23	57	26.8	532	3 US-09-218-550-6	Sequence 6, Appli
24	57	26.8	532	5 PCT-US92-01785-6	Sequence 6, Appli
25	57	26.8	532	5 PCT-US95-00454-6	Sequence 6, Appli
26	57	26.8	575	2 US-08-417-495-4	Sequence 0; Gaps
27	57	26.8	575	2 US-08-284-391B-4	Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
US-08-389-360-2  
; Sequence 2, Application US/08389360  
; Patent No. 5877017  
; GENERAL INFORMATION:  
; APPLICANT: van der Bruggen et al.  
; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/389,360  
; PUBLISH DATE: Herewith  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/196,630  
; FILING DATE: February 15, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/079,110  
; FILING DATE: June 17, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pascualina, Patricia A.  
; REGISTRATION NUMBER: 24,894  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 831-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 43  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-389-360-2

Query Match Best Local Similarity 100.0%; Score 213, DB 2; Length 43;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAASQIQLQARLMKEESPVVSVRLEPDGTALCFIF 43  
 Db 1 MAARAVFLAASQIQLQARLMKEESPVVSVRLEPDGTALCFIF 43

## RESULT 2

US-09-038-328-2  
 ; Sequence 2, Application US/09038328

; Patent No. 6110694

GENERAL INFORMATION:

APPLICANT: van der Bruggen et al.  
 TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch  
 STREET: 805 Third Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/435 524  
 FILING DATE: 08/No. 6465184-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/038,328

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/079,110

FILING DATE: June 17, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Pasqualini, Patricia A.

REGISTRATION NUMBER: 34,894

REFERENCE/DOCKET NUMBER: LUD 5310.2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 43

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-435-524-2

Query Match 100.0%; Score 213; DB 4; Length 43;

Best Local Similarity 100.0%; Pred. No. 1.e-26;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAASQIQLQARLMKEESPVVSVRLEPDGTALCFIF 43

Db 1 MAARAVFLAASQIQLQARLMKEESPVVSVRLEPDGTALCFIF 43

## RESULT 4

US-09-382-497-2

Sequence 2, Application US/09382497

; Patent No. 6638512

GENERAL INFORMATION:

APPLICANT: van der Bruggen et al.

TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/382,497

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/389,360

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/389,360

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/389,360

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/382,497

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/389,360

CLASSIFICATION:

FILING DATE:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/079,110  
 FILING DATE: June 17, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,894  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 338-3884  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 43  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: Linear  
 US-09-312-497-2

Query Match Best Local Similarity 100.0%; Pred. No. 8.5e-09; Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAQLQQLQARLME 22  
 Db 1 MAARAVFLAQLQQLQARLME 22

RESULT 6  
 US-08-196-630A-9  
 Sequence 9, Application US/08196630A  
 ; GENERAL INFORMATION:  
 ; PATENT NO. 568386  
 ; APPLICANT: van der Bruggen, Pierre  
 ; ADDRESS: Boon-Falleur, Thierry  
 ; TITLE OF INVENTION: ISOLATED PEPTIDES WHICH FORM COMPLEXES WITH MHC MOLECULE HLA-C-CLONE 10 AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York City  
 ; STATE: New York  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION NUMBER: US/08/196,630A  
 ; APPLICATION NUMBER: 08/079,110  
 ; FILING DATE: 15-FEB-1994  
 ; CLASSIFICATION: 436  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/079,110  
 ; FILING DATE: 17-JUN-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hanson, No. 568386man D.  
 ; REGISTRATION NUMBER: 30,946  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; TELEFAX: (212) 838-3884  
 ; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 16 amino acid residues  
 ; TYPE: amino acid  
 ; TOPOLOGY: Linear  
 US-08-196-630A-9.

Query Match Best Local Similarity 100.0%; Pred. No. 0.00016; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAQLQQLQARLME 16  
 Db 1 MAARAVFLAQLQQLQARLME 16

RESULT 7  
 US-09-621-976-5718  
 Sequence 5718, Application US/09621976  
 ; GENERAL INFORMATION:  
 ; PATENT NO. 6639063  
 ; APPLICANT: Giordano, J.Y.  
 ; ADDRESS: Dumas Milne Edwards, J.B.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET\_054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21

Query Match 46.0%; Score 98; DB 1; Length 22;



RESULT 11  
US-09-970-516-2  
ORGANISM: Homo sapiens

Query Match Similarity 29.1%; Score 62; DB 4; Length 384;  
Best Local Similarity 40.0%; Pred. No. 0.19; Indels 6; Gaps 1;  
Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

3 ARAVFLASQQLQARLIMKEES -----PVVSVRLEPEDG 36  
295 SRAMLLRFLAMEKGRAEMEYCPYLVTVPPVAAFRLEPKDG 334

RESULT 10  
US-08-284-391B-29  
Sequence 29, Application US/08284391B  
Patent No. 5,851,928  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
ADDRESS: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/218,950  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
REFERENCE/POCKET NUMBER: 00786/247001  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-284-391B-29

RESULT 12  
US-08-867-149-1  
Sequence 1, Application US/08867149  
Patent No. 5,921,76  
GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: Antibody Against a Host Cell  
TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure

Query Match Similarity 26.8%; Score 57; DB 2; Length 398;  
Best Local Similarity 35.1%; Pred. No. 1.3; Indels 2; Gaps 1;  
Matches 13; Conservative 7; Mismatches 15; Indels 2; Gaps 1;

6 VFLASQQLQARLIMKEESPVVSVRLEPEDGTAFCI 42  
338 LMISLKLKENKEAVSKREKPV -WVNPPAGWQCLL 372  
Qy Ddb

NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Maria C.H. Lin  
 STREET: 345 Park Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10154-0053  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version  
 SOFTWARE: #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/867,149  
 FILING DATE: 28-Feb-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/657,149  
 FILING DATE: 03-June-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maria C.H. Lin  
 REGISTRATION NUMBER: 29,323  
 REFERENCE/DOCKET NUMBER: 1151-4145  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 751-6849  
 TELEFAX: (212) 415-8745  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 433 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-867-149-1

Query Match Score 57; DB 2; Length 433;  
 Best Local Similarity 35.1%; Pred. No. 1.4;  
 Matches 13; Conservative 7; Mismatches 15; Indels 2; Gaps 1;

Qy 6 VPIALSAQLQARLMKKEESPVVSVRLEPDGTALCFI 42  
 Db 313 LMLSLKLENKEAKVSKREKEVW-VWLNPBAGMNOCLL 347

RESULT 14  
 US-09-100-409A-1  
 Sequence 1, Application US/09100409A  
 Patent No. 6093088  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Chang Yi  
 TITLE OF INVENTION: PEPTIDE COMPOSITION FOR  
 PREVENTION AND TREATMENT OF HIV INFECTION AND  
 IMMUNE DISORDERS  
 TITLE OF INVENTION: IMMUNE DISORDERS  
 NUMBER OF SEQUENCES: 64  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORGAN & FINNEGAN  
 STREET: 345 Park Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10154-0054  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version  
 SOFTWARE: #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/100,409A  
 FILING DATE:  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME:  
 REGISTRATION NUMBER:  
 REFERENCE/DOCKET NUMBER: 1151-4154  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-751-6849  
 TELEFAX: 212-751-6849  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 433 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-100-409A-1

Query Match Score 57; DB 3; Length 433;  
 Best Local Similarity 35.1%; Pred. No. 1.4;  
 Matches 13; Conservative 7; Mismatches 15; Indels 2; Gaps 1;

Qy 6 VPIALSAQLQARLMKKEESPVVSVRLEPDGTALCFI 42  
 Db 313 LMLSLKLENKEAKVSKREKEVW-VWLNPBAGMNOCLL 347

RESULT 15  
 US-08-374-1  
 Sequence 1, Application US/08808374  
 Patent No. 5961976  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Chang Yi  
 TITLE OF INVENTION: Antibody Against a Host Cell  
 TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure  
 TITLE OF INVENTION: Protection from Infection by HIV Primary Isolates  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Maria C.H. Lin  
 STREET: 345 Park Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10154-0053  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version  
 SOFTWARE: #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/808,374  
 FILING DATE: 28-Feb-1997  
 CLASSIFICATION: 435

Qy 6 VFLALSAQLQARLMKEESPVVSRLPEDGTALCFI 42  
 Db : |:| : |:| : |:| : |:| : |:| : |:| : |:| : |:|  
 313 LMLSLKLENKEAKVSKREKEV- -WVLNPEAGMnQCLL 347

## RESULT 15

US-08-466-368-4

Sequence 4, Application US/08466368

GENERAL INFORMATION:

APPLICANT: Maddon, Paul J.

APPLICANT: Littman, Dan R.

APPLICANT: Chess, Leonard

APPLICANT: Axel, Richard

APPLICANT: Weis, Robin

APPLICANT: McDougal, J. S.

TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN

TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper S. Dunham LLP

STREET: 1185 Avenue of Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.3.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,368

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 24577-E1-B/JPW/AKC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 458 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-466-368-4

Query Match 26.8%; Score 57; DB 3; Length 458;  
 Best Local Similarity 35.1%; Pred. No. 1.5%;  
 Matches 13; Conservative 7; Mismatches 15; Indels 2; Gaps 1;Qy 6 VFLALSAQLQARLMKEESPVVSRLPEDGTALCFI 42  
 Db : |:| : |:| : |:| : |:| : |:| : |:| : |:|  
 338 LMLSLKLENKEAKVSKREKEV- -WVLNPEAGMnQCLL 372

Search completed: July 7, 2004, 17:10:12

Job time : 28 secs

**This Page Blank (uspto)**

Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: July 7, 2004, 17:09:14 ; Search time 46 Seconds  
(without alignments)  
290.983 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213

Sequence: 1 MARAVFLAQLQQARLM.....SPVVSRLPEFDTALCFIF 43

Scoring table: BLOSUM62

Gapp 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
 1: /cgmn\_6/ptodata/1/pubpaas/US07\_PUBCOMB.pep:  
 2: /cgmn\_6/ptodata/1/pubpaas/PCT\_NEW\_PUB\_pep:  
 3: /cgmn\_6/ptodata/1/pubpaas/US06\_PUB\_pep:  
 4: /cgmn\_6/ptodata/1/pubpaas/US05\_PUBCOMB.pep:  
 5: /cgmn\_6/ptodata/1/pubpaas/US05\_PUB\_pep:  
 6: /cgmn\_6/ptodata/1/pubpaas/PCRTS\_NEW\_PUB\_pep:  
 7: /cgmn\_6/ptodata/1/pubpaas/US08\_PUB\_pep:  
 8: /cgmn\_6/ptodata/1/pubpaas/US08\_PUBCOMB.pep:  
 9: /cgmn\_6/ptodata/1/pubpaas/US09\_PUBCOMB.pep:  
 10: /cgmn\_6/ptodata/1/pubpaas/US09\_PUBCOMB.pep:  
 11: /cgmn\_6/ptodata/1/pubpaas/US09C\_PUBCOMB.pep:  
 12: /cgmn\_6/ptodata/1/pubpaas/US09c\_PUB\_pep:  
 13: /cgmn\_6/ptodata/1/pubpaas/US09c\_PUBCOMB.pep:  
 14: /cgmn\_6/ptodata/1/pubpaas/US10\_PUBCOMB.pep:  
 15: /cgmn\_6/ptodata/1/pubpaas/US10c\_PUBCOMB.pep:  
 16: /cgmn\_6/ptodata/1/pubpaas/US10c\_PUB\_pep:  
 17: /cgmn\_6/ptodata/1/pubpaas/US06\_PUB\_pep:  
 18: /cgmn\_6/ptodata/1/pubpaas/US06\_PUBCOMB.pep:  
 RESULT 1  
 US-10-296-734-825 ; Sequence 825, Application US/10296734  
 ; Publication No. US20040054137A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thompson, Scott A  
 ; TITLE OF INVENTION: Synthetic molecules and uses therefor  
 ; FILE REFERENCE: Savine  
 ; CURRENT APPLICATION NUMBER: US10/296,734  
 ; CURRENT FILING DATE: 2003-08-04  
 ; PRIOR APPLICATION NUMBER: AU PQ7761/00  
 ; PRIOR FILING DATE: 2000-05-26  
 ; NUMBER OF SEQ ID NOS: 1507  
 ; SEQ ID NO: 825  
 ; SOFTWARE: PatentIn version 3.2  
 ; LENGTH: 43  
 ; TYPE: PRT  
 ; FEATURE:  
 ; ORGANISM: Artificial  
 ; OTHER INFORMATION: BAGE consensus polypeptide  
 ; US-10-296-734-825

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	43	12	US-10-296-734-825
2	213	100.0	43	14	US-10-207-655-81
3	213	100.0	43	14	US-10-081-108-2
4	156	73.2	30	12	US-10-296-734-1214
5	156	73.2	3541	12	US-10-296-734-1454
6	126	59.2	30	12	US-10-296-734-1212
7	87	40.8	17	12	US-10-296-734-1216
8	62	29.1	293	10	US-09-933-767-328
9	62	29.1	293	12	US-10-004-860-128
10	62	29.1	293	14	US-10-023-262-328
11	62	29.1	333	15	US-10-264-237-585
12	62	29.1	368	14	US-10-053-510-21
13	62	29.1	368	15	US-10-348-052-21
14	62	29.1	384	9	US-09-784-810CA-2
15	62	29.1	384	9	US-09-970-516-2

RESULT 2  
 US-10-207-655-81 ; Sequence 81, Application US/10207655  
 ; Publication No. US2003011859A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ledbetter, Jeffrey A.

APPLICANT: Hayden-Ledbetter, Martha S.  
 TITLE OF INVENTION: BOUNDARY DOMAIN IMMUNOGLOBULIN FUSION PROTEINS  
 FILE REFERENCE: 39069.401C1  
 CURRENT APPLICATION NUMBER: US/10/207,655  
 CURRENT FILING DATE: 2003-07-25  
 NUMBER OF SEQ ID NOS: 426  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO: 81  
 LENGTH: 43  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-207-655-81

Query Match 100.0%; Score 213; DB 14; Length 43;  
 Best Local Similarity 100.0%; Prod. No. 6.7e-23;  
 Matches 43; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 1 MAARAVFLAQLQQLMKKEESPVVSRLEPDGTALCFIF 43  
 Db 1 MAARAVFLAQLQQLMKKEESPVVSRLEPDGTALCFIF 43

RESULT 4  
 US-10-296-734-1214  
 ; Sequence 1214, Application US/10296734  
 ; Publication No. US20040054137A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thompson, Scott A.  
 ; TITLE OF INVENTION: Synthetic molecules and uses therefor  
 ; FILE REFERENCE: Savine  
 ; CURRENT APPLICATION NUMBER: US/10/296,734  
 ; CURRENT FILING DATE: 2003-08-04  
 ; PRIOR APPLICATION NUMBER: AU P27761/00  
 ; PRIOR FILING DATE: 2000-05-26  
 ; NUMBER OF SEQ ID NOS: 1507  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO: 1214  
 ; LENGTH: 30  
 ; TYPE: PRT  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: BAGE segment 2

Query Match 73.2%; Score 156; DB 12;  
 Best Local Similarity 100.0%; Prod. No. 6.3e-15;  
 Matches 30; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 14 LIQARLMKKEESPVVSRLEPDGTALCFIF 43  
 Db 1 LIQARLMKKEESPVVSRLEPDGTALCFIF 30

RESULT 5  
 US-10-296-734-1454  
 ; Sequence 1454, Application US/10296734  
 ; Publication No. US20040054137A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thompson, Scott A.  
 ; TITLE OF INVENTION: Synthetic molecules and uses therefor  
 ; FILE REFERENCE: Savine  
 ; CURRENT APPLICATION NUMBER: US/10/296,734  
 ; CURRENT FILING DATE: 2003-08-04  
 ; PRIOR APPLICATION NUMBER: AU P27761/00  
 ; PRIOR FILING DATE: 2000-05-26  
 ; NUMBER OF SEQ ID NOS: 1507  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO: 1454  
 ; LENGTH: 3541  
 ; TYPE: PRT  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: Melanoma cancer specific savine

Query Match 73.2%; Score 156; DB 12;  
 Best Local Similarity 100.0%; Prod. No. 1.7e-12;  
 Matches 30; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 14 LIQARLMKKEESPVVSRLEPDGTALCFIF 43  
 Db 2777 LIQARLMKKEESPVVSRLEPDGTALCFIF 2806

RESULT 6  
 US-10-296-734-1212  
 ; Sequence 1212, Application US/10296734  
 ; Publication No. US20040054137A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thompson, Scott A.  
 ; TITLE OF INVENTION: Synthetic molecules and uses therefor  
 ; FILE REFERENCE: Savine  
 ; CURRENT APPLICATION NUMBER: US/10/296,734  
 ; CURRENT FILING DATE: 2003-08-04  
 ; PRIOR APPLICATION NUMBER: AU P27761/00  
 ; PRIOR FILING DATE: 2000-05-26  
 ; NUMBER OF SEQ ID NOS: 1507  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO: 1212  
 ; LENGTH: 3541  
 ; TYPE: PRT  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: Melanoma cancer specific savine

Query Match 100.0%; Score 213; DB 14;  
 Best Local Similarity 100.0%; Prod. No. 6.7e-23;

US-10-081-108-2

Sequence 1212, Application US/10296734  
 Publication No. US2004054137A1  
 GENERAL INFORMATION:  
 APPLICANT: Thompson, Scott A  
 TITLE OF INVENTION: Synthetic molecules and uses therefor  
 FILE REFERENCE: Savine  
 CURRENT APPLICATION NUMBER: US/10/296,734  
 PRIOR FILING DATE: 2003-08-04  
 PRIOR APPLICATION NUMBER: AU PQ7761/00  
 NUMBER OF SEQ ID NOS: 1507  
 SEQ ID NO: 1212  
 LENGTH: 30  
 TYPE: PRT  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO: 1212  
 LENGTH: 30  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: BAGE segment 1  
 US-10-296-734-1212

Query Match 59.2%; Score 126; DB 12; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLASQQLQARLAKKEESPVVS 28  
 Db 3 MAARAVFLASQQLQARLAKKEESPVVS 30

RESULT 7  
 US-10-296-734-1212  
 Sequence 1211, Application US/10296734  
 Publication No. US2004054137A1  
 GENERAL INFORMATION:  
 APPLICANT: Thompson, Scott A  
 TITLE OF INVENTION: Synthetic molecules and uses therefor  
 FILE REFERENCE: Savine  
 CURRENT FILING DATE: 2003-08-04  
 PRIOR APPLICATION NUMBER: AU PQ7761/00  
 PRIOR FILING DATE: 2003-05-26  
 NUMBER OF SEQ ID NOS: 1507  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO: 1216  
 LENGTH: 17  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: BAGE segment 3  
 US-10-296-734-1216

Query Match 40.8%; Score 87; DB 12; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 WRLEPEDGTALCFIF 43  
 Db 1 WRLEPEDGTALCFIF 15

RESULT 8  
 US-09-933-767-328  
 Sequence 328, Application US/099333767  
 Publication No. US20030181692A1  
 GENERAL INFORMATION:  
 APPLICANT: Ni et al.  
 TITLE OF INVENTION: 207 Human Secreted Proteins  
 FILE REFERENCE: P200782  
 CURRENT APPLICATION NUMBER: US/09/933,767  
 CURRENT FILING DATE: 2001-08-22  
 PRIOR APPLICATION NUMBER: PCT/US01/05614  
 PRIOR FILING DATE: 1997-06-06



; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,916  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,373  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,875  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,374  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,917  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,949  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,883  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,897  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,898  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,962  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,963  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,877  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,878  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/070,923  
 ; EARLIER FILING DATE: 1997-12-18  
 ; EARLIER APPLICATION NUMBER: 60/092,921  
 ; EARLIER FILING DATE: 1998-07-15  
 ; EARLIER APPLICATION NUMBER: 60/094,657  
 ; EARLIER FILING DATE: 1998-07-30  
 ; NUMBER OF SEQ ID NOS: 1227  
 ; SEQ ID NO: 328  
 ; LENGTH: 293  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (36)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-10-023-282-328

Query Match 29.1%; Score 62; DB 14; Length 293;  
 Best Local Similarity 40.0%; Pred. No. 2.5;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLASQQLQARLMKES----PVYWSWRLEPEDG 36  
 Db 204 SRAMLLRFLAMEKGRRHMEYCPCPYVPPVAAFRLEPKDG 243

RESULT 11  
 US-10-264-237-2585  
 ; Sequence 2585, Application US/10264237  
 ; GENERAL INFORMATION:  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; CURRENT FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: PCT/US01/16450  
 ; PRIOR FILING DATE: 2001-05-18  
 ; PRIOR APPLICATION NUMBER: US 60/205,515  
 ; PRIOR FILING DATE: 2000-05-19  
 ; NUMBER OF SEQ ID NOS: 2876  
 ; SOFTWARE: PatentIn Ver. 3.1  
 ; SEQ ID NO: 2585

Query Match 29.1%; Score 62; DB 15; Length 368;  
 Best Local Similarity 40.0%; Pred. No. 3.3;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLASQQLQARLMKES----PVYWSWRLEPEDG 36  
 Db 279 SRAMLLRFLAMEKGRRHMEYCPCPYVPPVAAFRLEPKDG 318

RESULT 13  
 US-10-348-052-21  
 ; Sequence 21, Application US/10348052  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saba, Julie D.  
 ; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND  
 ; METHODS OF USE THEREFOR  
 ; FILE REFERENCE: 200116-402C2  
 ; CURRENT APPLICATION NUMBER: US/10/053,510  
 ; CURRENT FILING DATE: 2002-01-17  
 ; NUMBER OF SEQ ID NOS: 21  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO: 21  
 ; LENGTH: 368  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-10-033-510-21

Query Match 29.1%; Score 62; DB 14; Length 368;  
 Best Local Similarity 40.0%; Pred. No. 3.3;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLASQQLQARLMKES----PVYWSWRLEPEDG 36  
 Db 279 SRAMLLRFLAMEKGRRHMEYCPCPYVPPVAAFRLEPKDG 318

RESULT 15  
 US-10-348-052-21  
 ; Sequence 21, Application US/10348052  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saba, Julie D.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION  
 ; TITLE OF INVENTION: OF SPHINGOLIPID METABOLISM AND/OR SIGNALING  
 ; FILE REFERENCE: 200116-405  
 ; CURRENT APPLICATION NUMBER: US/10/348,052  
 ; CURRENT FILING DATE: 2003-01-17  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO: 21  
 ; LENGTH: 368  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-10-348-052-21

Query Match 29.1%; Score 62; DB 15; Length 368;  
 Best Local Similarity 40.0%; Pred. No. 3.3;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLASQQLQARLMKES----PVYWSWRLEPEDG 36  
 Db 279 SRAMLLRFLAMEKGRRHMEYCPCPYVPPVAAFRLEPKDG 318

RESULT 14  
IS-09-784-810A-2  
Sequence 2, Application US/09784810A  
Patent No. US20020082203A1  
GENERAL INFORMATION:  
APPLICANT: RASTELLI, LUCA  
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
TITLE OF INVENTION: SAME  
FILE REFERENCE: 10716-08  
CURRENT APPLICATION NUMBER: US/09/784.810A  
CURRENT FILING DATE: 2001-02-14  
PRIOR APPLICATION NUMBER: 60/182,360  
PRIOR FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: 60/191,261  
PRIOR FILING DATE: 2000-03-22  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 384  
TYPE: PRT  
ORGANISM: Homo sapiens  
IS-09-784-810A-2

Query	Match      29.1% ; Best Local Similarity    40.0% ; Matches    16;	Score 62;      DB 9;      Length 384; Pred. No. 3.4; Mismatches 8;	Indels 10;	Gaps 6;	1;
Sbj	<pre> 3 ARAVFLAQLQARLIMKES-----PQVSMLPEPDG 36 : : : : : : : : : : : : : : : : : : :  295 SRAMMIRFLAMEKGRRHMEYECPVLYVVPVAFLERPKDG 334 </pre>				

REBUST 15  
US-09-970-516-2  
Sequence 2, Application US/09970516  
Patent No. US20020099029A1  
GENERAL INFORMATION:  
APPLICANT: No. US20020099029A1  
TITLE OF INVENTION: Induction of blood vessel formation through administration of  
TITLE OF INVENTION: Polynucleotides encoding sphingosine kinases  
FILE REFERENCE: 4-31617  
CURRENT APPLICATION NUMBER: US/09/970,516  
CURRENT FILING DATE: 2001-10-04  
NUMBER OF SEQ ID NS: 6  
SOFTWARE: PatentIn version 3.1

Search completed: July 7, 2004, 17:14:58  
Job time : 48 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 7, 2004, 17:05:13 ; Search time 20 Seconds

(without alignments)  
 206.812 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213

Sequence: 1 MARAVFIALLASQILQARLM.....SPVVSRLLEPEDGTALCFIF 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR78:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	55.5	26.1	359	2	JC7280	cytokine receptor-transposse - Esch
2	55	25.8	154	2	F69520	probable excrinucle
3	54.5	25.6	994	1	TOBCTS	hypothetical prote
4	54	25.4	295	2	G90934	conserved hypothet
5	54	25.4	295	2	C85783	unknown protein [i]
6	54	25.4	302	2	AD0709	structure homolog
7	54	25.4	552	2	T00459	Ybd Protein - Esc
8	54	25.4	552	2	AC0609	cdd3-like protein
9	54	25.4	552	2	A96802	probable thiosulfa
10	54	25.4	699	2	E97591	probable membrane
11	53.1	155	2	D64826	hypothetical prote	
12	53	24.9	552	2	A64714	helicase - Helicob
13	53	24.9	552	2	F85599	hypothetical prote
14	53	24.9	552	2	T04489	protein kinase hom
15	52.5	24.6	243	2	T51377	hypothetical prote
16	52	24.4	346	2	T05639	ATP-stimulated glu
17	52	24.4	445	2	A64714	probable ATP-depen
18	52	24.4	945	2	JN0606	fibrinogen-binding
19	51	23.9	254	2	T09009	hypothetical prote
20	51	23.9	284	2	G70732	gp1 protein - Myco
21	51	23.9	332	2	A99738	
22	51	23.9	332	2	B85588	
23	51	23.9	332	2	C64816	
24	51	23.9	524	1	S36175	
25	51	23.9	524	2	A71805	
26	51	23.9	946	2	AC1214	
27	50.5	23.7	862	2	S13979	
28	50	23.5	58	2	A72800	
29	50	23.5	270	2		

30	50	23.5	334	2	G75344	probable polyferre
31	50	23.5	367	2	AI2606	hypotheical prote
32	50	23.5	367	2	H97388	periplasmic mann
33	50	23.5	396	2	AH2275	two-component resp
34	50	23.5	503	2	T48825	hypothetical prote
35	49.5	23.2	151	2	AB0331	sigma B factor reg
36	49.5	23.2	247	2	A43846	lipoprotein LPPA -
37	49.5	23.2	254	2	C42691	fibroblast growth
38	49.5	23.2	769	2	S16236	fibroblast growth
39	49.5	23.2	821	1	TVHF02	fibroblast growth
40	49.5	23.2	822	2	A45081	keratinocyte growt
41	49.5	23.2	822	2	A41194	extragenic suppres
42	49	23.0	267	2	S74899	hypothetical prote
43	49	23.0	319	2	T27862	hypothetical prote
44	49	23.0	353	2	B72468	flavohemoglobin hm
45	49	23.0	411	2	B83782	

## ALIGNMENTS

RESULT 1	JC7280	cytokine receptor-like molecule-2 - mouse
	C;Species: Mus musculus (house mouse)	
	C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000	
	C;Accession: JC7280	
	R;Hiroyama, T.; Iwama, A.; Morita, Y.; Nakamura, Y.; Shibuya, A.; Nakauchi, H.	
	Biochem Biophys Res Commun. 272, 224-229, 2000	
	A;Title: Molecular cloning and characterization of CRLM-2, a novel type I cytokine rece	
	A;Reference number: JC7280	
	A;Contents: Embryo	
	A;Accession: JC7280	
	A;Molecule type: mRNA	
	A;Residues: 1-359 <HR>	
	A;Cross references: DDBJ:AB039945	
	C;Genetics:	
	A;Gene: crlm-2	
	C;Keywords: cytokine; embryo; receptor; signal transduction; transmembrane protein	
Query	KE3SPVV--SRLPEPDGTALCFI 42	Score 55.5; DB 2; Length 359;
Match	21 KE3SPVV--SRLPEPDGTALCFI 42	Best Local Similarity 44.0%; Pred. No. 9.6;
Match	11; Conservative	Matches 7; Mismatches 4; Indels 3; Gaps 1;
Db	301 EBDIDLHTKAYEPBEGTSCTV 325	
RESULT 2	F65520	
	conserved hypothetical protein AF2166 - Archaeoglobus fulgidus	
	C;Species: Archaeoglobus fulgidus	
	C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997	
	R;Klenk, H.P.; Clayton, R.A.; Tomb, J.-F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso	
	; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.	
	Glodner, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; McDonald, L.	
	Nature 390, 364-370, 1997	
	A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Kaine, B.P.; Sykes, S	
	Smith, H.O.; Woese, C.R.; Venter, J.C.	
	A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae	
	A;Reference number: A69250; MNUID:98049343; PMID:9389475	
	A;Accession: F65520	
	A;Status: preliminary; nucleic acid sequence not shown; translation not shown	
	A;Molecule type: DNA	
	A;Residues: 1-154 <KLE>	
	A;Cross references: GB:AE000955; GB:AE000782; PID:92689278; PMID:9268909..1; PMID:926483	
Query	Match 25.8%; Score 55; DB 2; Length 154;	
Match	Best Local Similarity 34.2%; Pred. No. 4.6%; Mismatches 10; Indels 15; Gaps 0;	
Matches	13; Conservative	

probable excinuclease subunit 22771 [imported] - Escherichia coli (strain O157:H7, substrate) Qy 2 AARAVFLAALSQQLQRMLKEESPVWSWRLEPEDDTAL 39 Db 24 ATNAIFFLAYALITKSKVDXKIKFPVITMILVPLDFAAM 61

**RESULT 3**  
Tn2501 - Escherichia coli transposon Tn2501  
C;Species: Escherichia coli  
C;Accession: 30-Jun-1988 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C;Accession: S06302; A27758  
R;Turner, A.K.; Grinsted, J.  
R;Michiels, T.; Cornelis, G.; Elkins, K.; Grinsted, J.  
J.; Bacteriol., 169, 624-631, 1987  
A;Title: Tn2501, a component of the lactose transposon Tn951, is an example of a new cat  
A;Reference number: A91859; PMID:88096496; PMID:28277105  
A;Accession: S06302  
A;Molecule type: DNA  
A;Residues: 1-94 <TUR>  
A;Cross-references: EMBL:Y00502; NID:943098; PIDN:CA68555.1; PID:943099  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res., 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A;Reference number: A99629; PMID:21156231; PMID:11258796

probable excinuclease subunit 22771 [imported] - Escherichia coli (strain O157:H7, substrate) Qy 2 AARAVFLAALSQQLQRMLKEESPVWSWRLEPEDDTAL 39 Db 24 ATNAIFFLAYALITKSKVDXKIKFPVITMILVPLDFAAM 61

**RESULT 4**  
Tn2501 - Escherichia coli transposon Tn2501  
C;Species: Escherichia coli  
C;Accession: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res., 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A;Reference number: A99629; PMID:21156231; PMID:11258796

probable excinuclease subunit [imported] - Escherichia coli (strain O157:H7, substrate) Qy 2 AARAVFLAALSQQLQRMLKEESPVWSWRLEPEDDTAL 39 Db 24 ATNAIFFLAYALITKSKVDXKIKFPVITMILVPLDFAAM 61

**RESULT 5**  
Tn2501 - Escherichia coli transposon Tn2501  
C;Species: Escherichia coli  
C;Accession: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; PMID:21074935; PMID:11206551  
A;Status: preliminary  
A;Accession: C85783  
A;Molecule type: DNA  
A;Residues: 1-295 <STO>  
A;Cross-references: GB:AB005174; NID:91251560; PIDN:AAG56727.1; GSDB:GN00145; PMID:11206551  
A;Experimental source: strain O157:H7, substring EDL933  
C;Genetics:  
A;Gene: Z2771

Query Match 25.4%; Score 54; DB 2; Length 295;  
Best Local Similarity 50.0%; Pred. No. 13; Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 10 LSAQLIQLKMEESPVWSWRL 31  
Db 89 ICALLIEEARLILKEQQPLFNKRL 110

**RESULT 6**  
E64933 - hypothetical protein b1741 - Escherichia coli (strain K-12)  
C;Species: Escherichia coli  
C;Accession: 12-Sep-1997 #sequence\_revision 17-Sep-1999 #text\_change 01-Mar-2002  
R;Blattner, F.R.; Plunkett III, G.; Perna, N.T.; Burland, V.; Riley, M.; Cole, R.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64730; PMID:9742667; PMID:978503  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-295 <BLAT>  
A;Cross-references: GB:AB000269; GB:J00096; NID:91788033; PIDN:AACT74811.1; PMID:91788037; A;Experimental source: strain K-12, substring MGJ655

Query Match 25.4%; Score 54; DB 2; Length 295;  
Best Local Similarity 50.0%; Pred. No. 13; Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 10 LSAQLIQLKMEESPVWSWRL 31  
Db 89 ICALLIEEARLILKEQQPLFNKRL 110

**RESULT 7**  
AD0709 - conserved hypothetical protein STY1804 [imported] - Salmonella enterica subsp. enterica C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called *Salmonella typhi*  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AD0709  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, S.; T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moul, S.; O'Gara, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar  
A;Reference number: AD0502; PMID:2453497; PMID:11677608  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-302 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD02045.1; PMID:916502882; GSDB:GN00176 C85783

C;Genetics: SRY1804

A;Gene: A96802

Query Match Score 54; DB 2; Length 302;

Best Local Similarity 50.0%; Pred. No. 13; Mismatches 6; Indels 0; Gaps 0;

Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 10 LSAQLQIARLMKEESPVWSWPL 31

Db 98 MGALIIEARVLIKEQQPFLENFL 119

RESULT 8 hypothetical protein T14N5.16 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 01-Feb-1993 #sequence\_revision 01-Feb-1993 #text\_change 22-Oct-1999

R;Federer, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo,

;Vysotskaya, V.S.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.

A;Experimental source: submitted to the EMBL Data Library, September 1998

A;Reference number: Z14152

A;Accession: T00459

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-525 &lt;SPED&gt;

A;Cross-references: EMBL:AC004260; PID:93176694; PID:93540209; GSPDB:GN00059; ATSP:T14N5

A;Gene: ATSP:T14N5.16

A;Map position: 1

A;Introns: 83/3; 112/3; 143/3; 169/2; 212/3; 335/3; 363/1; 400/3; 429/3; 471/3

Qy 25.4%; Score 54; DB 2; Length 525;

Best Local Similarity 40.0%; Pred. No. 23; Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 6 VFIAASQIQLQIARLMKEESPVWSWR 30

Qy 25 ILTAGDRKLLKVLLKEETIVSWK 49

RESULT 9 conserved hypothetical protein STY0935 [imported] - Salmonella enterica subsp. enterica

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: This species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AC0609

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-552 &lt;PAR&gt;

A;Cross-references: GB:AL513382; PID:CA05341.1; PID:g16502105; GSEDB:GN00176

A;Gene: STY0935

Query Match Score 54; DB 2; Length 552;

Best Local Similarity 34.9%; Pred. No. 24; Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

Db 11 SAQLIQ-----ARLMKEESPVWSWR----LEPEDGTAFC 41

Qy 324 SGELLSTPVEHVRVRERESSAAWLGTCGLSAEDGRRIAF 366

Db 11 AVILARM--ESPLAGWQLEP 29

RESULT 10 unknown protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C;Accession: A96802

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

Ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A96802

A;Cross-references: &lt;STO&gt;

A;Residues: 1-699

C;Genetics:

A;Map position: 1

A;Gene: F2P24.2

A;Map type:保守

Query Match Score 54; DB 2; Length 699;

Best Local Similarity 40.0%; Pred. No. 31; Matches 10; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Qy 6 VFIAASQIQLQIARLMKEESPVWSWR 30

Db 25 ILTAGDRKLLKVLLKEETIVSWK 49

RESULT 11 structure homolog of thermus thermophilus 30S ribosomal chain M [imported] - Agrobacter

C;Species: Agrobacterium tumefaciens

C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 12-Jun-2003

C;Accession: E97591

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Ourollo, B.; Goldman

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 2320-2328, 2001

A;Title: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens

A;Reference number: A973559; MUID:21608551; PMID:11743194

A;Accession: E97591

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-155 &lt;KOR&gt;

A;Cross-references: GB:AE007869; PID:AAK87686.1; PID:915157045; GSPDB:GN00169

C;Genetics:

A;Map position: circular chromosome S13/S18

C;Superfamily: ribosomal protein S13/S18

Query Match Score 53.5%; DB 2; Length 155;

Best Local Similarity 54.5%; Pred. No. 74; Matches 12; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

Qy 12 AQIQLQIARLMKEESPVWSWR 33

Db 11 AVILARM--ESPLAGWQLEP 29

RESULT 12

D64826 ybID protein - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002

C;Accession: D64826

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference: A64720; MUID:7426617; PMID:9278503  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-552 <BLAT>  
A;Cross-references: GB:AE000189; GB:U0096; NID:91787097; PIDN: AAC73963.1; PID:91787102;  
A;Experimental source: strain K-12, substrate M91655  
C;Genetics:  
A;Gene: ybdJ  
C;Keywords: ATP; nucleotide binding; P-loop; transmembrane protein  
F;29-36/Region: nucleotide-binding motif A (P-loop)  
F;264-280/Domain: transmembrane #status predicted <IMM>

Query Match Score 24.9%; DB 2; Length 552;  
Best Local Similarity 34.9%; Pred. No. 33;  
Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

Qy 11 SAQLIQ----ARLMKERSPVWSWRLEP----EDGTAFCF 41  
Db 324 SGELSLSLTPVHVCRLVRESSRVAAWRLGPGLSTEDSRRI SF 366

RESULT 15

T04489 protein kinase homolog F8F16.40 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 30-Apr-1999  
C;Accession: T04489  
R;Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hoheisel, J.; Meier  
submitted to the Protein Sequence Database, April 1998  
A;Reference number: 215375  
A;Accession: T04489  
A;Molecule type: DNA  
A;Residues: 1-243 <BEV>  
A;Cross-references: EMBL:AL021633  
A;Experimental source: cultivar Columbia; BAC clone F8F16  
C;Genetics:  
A;Map Position: 4

B90749 Query Match Score 24.6%; DB 2; Length 243;  
A;Note: F8F16.40  
C;Species: Escherichia coli (strain O157:H7, substrate R1)  
C;Accession: B90749  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno  
A;Reference number: A99629; MUID:21156231; PMID:11255796  
A;Accession: B90749  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-552 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BAB34385.1; PID:G13360421; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrate RIMD 0509952  
C;Genetics:  
A;Gene: ECs0962

RESULT 13

B90749 Query Match Score 24.6%; DB 2; Length 243;  
Best Local Similarity 34.9%; Pred. No. 16;  
Matches 13; Conservative 7; Mismatches 5; Indels 5; Gaps 2;

Qy 15 LQARLMKE-DSPVWSW----RLEPDGTAFCF 42  
Db 24 LEDRLFQGDSPALSQTRRIAAEIGTVLFL 56

Search completed: July 7, 2004, 17:09:33  
Job time : 22 secs

F85599 Query Match Score 24.9%; DB 2; Length 552;  
Best Local Similarity 34.9%; Pred. No. 33;  
Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

Qy 11 SAQLIQ----ARLMKERSPVWSWRLEP----EDGTAFCF 41  
Db 324 SGELSLSLTPVHVCRLVRESSRVAAWRLGPGLSTEDSRRI SF 366

RESULT 14

F85599 hypothetical protein ybdJ [imported] - Escherichia coli (strain O157:H7, substrate EDL93  
C;Species: Escherichia coli (strain O157:H7, substrate EDL93)  
C;Accession: F85599  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lin, A.; Dimmler, E.; Potamousis, K.; Apodaca,  
Nature 401, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: F85599  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-552 <STO>  
A;Cross-references: GB:AE000174; NID:G1251906; PID:AA55258.1; GSPDB:GN00145; UWGP:Z11  
A;Experimental source: strain O157:H7, substrate EDL933  
C;Genetics:  
A;Gene: ybdJ

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 7, 2004, 16:58:33 ; Search time 12 Seconds  
(without alignments)  
186.385 Million cell updates/sec

Title: US-10-081-108-2  
Perfect score: 213  
Sequence: 1 MAARAYFLALSAQILQARLM.....SPVVSWRLEPEDGTLACFIF 43

## Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

## Total number of hits satisfying chosen parameters:

141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	213	100.0	43	1	BGE1_HUMAN	Q13072	homo sapien
2	206	96.7	43	1	BGE5_HUMAN	Q86927	homo sapien
3	177	83.1	109	1	BGE2_HUMAN	Q86730	homo sapien
4	177	83.1	109	1	BGE3_HUMAN	Q86729	homo sapien
5	173	81.2	39	1	BGE4_HUMAN	Q86928	homo sapien
6	62	29.1	384	1	SPH1_HUMAN	Q9yai1	homo sapien
7	55.5	26.1	359	1	CRL2_MOUSE	Q8C119	mus musculus
8	54.5	25.8	154	1	Y16_6_ARCFU	Q28116	archaeoglob
9	54.5	25.6	994	1	TNP5_ECOLI	P08514	escherichia
10	54	25.4	293	1	CHO_SALTY	Q8z695	salmonella
11	54	25.4	293	1	CHO_SALTY	Q8zpu5	salmonella
12	54	25.4	295	1	CHO_ECOLI	P59437	escherichia
13	54	25.4	295	1	CHO_ECOL6	P76213	escherichia
14	54	25.4	295	1	CHO_ECOLI	P59361	shigella f1
15	54	25.4	295	1	CHO_SHFL	P83038	lulus galli
16	54	25.4	1080	1	HDA4_CHICK	P256524	homo sapien
17	54	25.4	1084	1	HDA4_HUMAN	P75828	escherichia
18	53	24.9	552	1	YBJD_ECOLI	Q9mug8	mesotigma
19	52.5	24.6	170	1	YC66_MESVI	Q8Bf6	escherichia
20	52	24.4	332	1	YBHG_ECOL6	Q95950	mycobacterium
21	51	23.9	284	1	THT3_MYCTU	Q8xty9	escherichia
22	51	23.9	332	1	YBHG_ECOL7	P75777	escherichia
23	51	23.9	332	1	YBHG_ECOLI	Q833s16	shigella f1
24	51	23.9	332	1	YBHG_SHFL	Q29037	saimiri sci
25	51	23.9	457	1	CD4_SA1SC	P32189	homo sapien
26	51	23.9	524	1	GLPK_HUMAN	Q64516	mus musculus
27	51	23.9	524	1	GLPK_MOUSE	Q63060	rattus norvegicus
28	51	23.9	524	1	GLPK_RAT	Q922v6	mus musculus
29	51	23.9	1113	1	HDA5_BMD2	Q64197	mycobacterium
30	50	23.5	270	1	VGP1_BMD2	Q9ug10	homo sapien
31	50	23.5	708	1	TRAB_HUMAN	Q9ug16	homo sapien
32	50	23.5	1122	1	HDAS_HUMAN	Q9ukm7	homo sapien
33	49.5	23.2	699	1	MIB1_HUMAN		

## ALIGNMENTS

RESULT 1	BGE1_HUMAN	STANDARD;	PRT;	43 AA.
ID	BGE1_HUMAN			
AC	Q13072;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DB	B melanoma antigen 1 precursor (B melanoma antigen) (Antigen MZ2-BA).			
GN	BAGE OR BGE1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Buthidae; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TAXID	9606;			
[1]				
RN	SEQUENCE FROM N.A. (ISOFORM BAGE1A).			
RP	SEQUENCE FROM Melatoma;			
RC	TISSUE-Melatoma;			
RX	PubMed=895173;			
RA	Boel P., Waldmann C., Sensi M. I., Brasseur R., Renaud J.-C., Coulie P., Boon T., van der Bruggen P.;			
RA	"New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile."			
RT	"BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes."			
RT	Immunity 2:167-175(1995).			
[2]				
RN	ALTERNATIVE SPlicing.			
RP	MEDLINE=22349465; PubMed=12461691;			
RX	Ruan H., van der Bruggen P., Brun M.-E., Boyle S., Roizes G., De Sarro A.,			
RA	"New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile."			
RT	Eur. J. Hum. Genet. 10:833-840(2002).			
RA	FUNCTION:Unknown. Antigen recognized on a melanoma by autologous cytolytic T lymphocytes.			
[3]				
-1	-1 ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=1;			
CC	Comment=At least 6 different mRNAs, BAGE1a, BAGE1b, BAGE1c, BAGE1d, BAGE1e and BAGE1f, are produced by alternative splicing.			
CC	IsoID=C13072-1; Sequence=Displayed;			
CC	Name1; Synonyms=BAGE1a;			
CC	Sequence;			
CC	-1 TISSUE SPECIFICITY: Not expressed in normal tissues, except in testis. Expressed with significant proportion in melanoma, but also in tumors of various histological origins, such as bladder carcinomas, head and neck squamous cell carcinomas, lung and breast carcinomas. Not expressed in renal, colorectal and prostatic carcinomas, leukemias and lymphomas. More frequently expressed in metastatic melanomas than in primary melanomas.			
CC	-1 MISCELLANEOUS: The ancestral BAGE gene was generated by juxtapacentromeric reshuffling of the ML3 Gene. The BAGE family was expanded by juxtapacentromeric movement and/or acrocentric exchanges. BAGE family is composed of expressed genes that map to the juxtapacentromeric regions of chromosomes 13 and 21 and of unexpressed gene fragments that scattered in the juxtapacentromeric regions of several chromosomes, including chromosomes 9, 13, 18 and 21.			

CC -!- SIMILARITY: Belongs to the BAGE family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC use by non-profit institutions. There are no restrictions on its

CC modifiability and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; U19180; AAC50123.1; -.

DR EMBL; AAP0743.1; -.

DR EMBL; AF521550; AAO22670.1; -.

DR EMBL; AF521551; AAC012671.1; -.

DR EMBL; AF521552; AAC012672.1; -.

DR EMBL; AF521553; AAO22673.1; -.

DR EMBL; AF521554; AAC012674.1; -.

DR Genew; HGNC: 942; BAGE.

DR NM\_60517;

KW Antigen; Alternative splicing; Multigene family; Signal.

FT SIGNAL 17 POTENTIAL.

FT CHAIN 18 B MELANOMA ANTIGEN 1.

SQ SEQUENCE 43 AA; 4810 MW; 36F3BBCB4012F1BB CRC64;

Query Match 100.0%; Score: 213; DB 1; Length: 43;

Best Local Similarity 100.0%; Pred. No. 1..le-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAQLSAQQLQARLNKEESPVSVRLPEPDGTALCFIF 43

Db 1 MAARAVFLAQLSAQQLQARLNKEESPVSVRLPEPDGTALCFIF 43

RESULT 3

BGE2\_HUMAN STANDARD; PRT; 109 AA.

ID BGE2\_HUMAN

AC Q86Y30;

DT 10-OCT-2003 (Rel: 42, Created)

DT 10-OCT-2003 (Rel: 42, Last sequence update)

DT 10-OCT-2003 (Rel: 42, Last annotation update)

DE B melanoma antigen 2 precursor.

GN BAGE2.

RA Homo sapiens (Human)

RA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

RA Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

RA NCBITaxonID=9606;

RN [1] SEQUENCE FROM N.A.

RP TISSUE=Testis;

RC MEDLINE=22349415; PubMed=12451691;

RX Ruauult M., Van Der Bruggen P., Brun M.-E., Boyle S., Roizes G.,

RA Sario A.D.;

RA "New BAGE (B melanoma antigen) genes mapping to the juxtapacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile.",

RT Eur. J. Hum. Genet. 10:833-840(2002).

RN [2] SEQUENCE FROM N.A.

RP TISSUE=Testis;

RC MEDLINE=22563396; PubMed=12676563;

RX Ruauult M., Ventura M., Galtier N., Brun M.-E., Archidiacono N.,

RA Roizes G., De Sario A.;

RT "BAGE genes generated by juxtapacentromeric reshuffling in the hominidae lineage are under selective pressure.",

RA Genomics 81:391-399(2003).

CC -!- FUNCTION: Unknown. Candidate gene encoding tumor antigens.

CC -!- TISSUE SPECIFICITY: Not expressed in normal tissues except in testis. Expressed in melanoma, bladder and lung carcinomas.

CC -!- MISCELLANEOUS: The ancestral BAGE gene was generated by juxtapacentromeric reshuffling of the MLL3 gene. The BAGE family was expanded by juxtapacentromeric regions of chromosomes 13 and 21 and of exchanges. BAGE family is composed of expressed genes that map to the juxtapacentromeric regions of chromosomes 13 and 21 and of

CC regions of several chromosomes, including chromosomes 9, 13, 18 and 21.

CC -!- SIMILARITY: Belongs to the BAGE family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC use by non-profit institutions. There are no restrictions on its

CC modifiability and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; AF330516; AAO22636.1; -.

DR Genew; HGNC:15872; BAGES5

KW Antigen; Multigene family; Signal.

FT SIGNAL 17 POTENTIAL.

FT CHAIN 18 B MELANOMA ANTIGEN 5.

SQ SEQUENCE 43 AA; 4710 MW; C6F3AACB4012F1A6 CRC64;

Query Match 96.7%; Score: 206; DB 1; Length: 43;

Best Local Similarity 97.7%; Pred. No. 1..e-21;

Matches 42; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAARAVFLAQLSAQQLQARLNKEESPVSVRLPEPDGTALCFIF 43

Db 1 MAAGAVFLAQLSAQQLQARLNKEESPVSVRLPEPDGTALCFIF 43

RESULT 3

BGE2\_HUMAN STANDARD; PRT; 109 AA.

ID BGE2\_HUMAN

AC Q86Y30;

DT 10-OCT-2003 (Rel: 42, Created)

DT 10-OCT-2003 (Rel: 42, Last sequence update)

DT 10-OCT-2003 (Rel: 42, Last annotation update)

DE B melanoma antigen 2 precursor.

GN BAGE2.

RA Homo sapiens (Human)

RA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

RA Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

RA NCBITaxonID=9606;

RN [1] SEQUENCE FROM N.A.

RP TISSUE=Testis;

RC MEDLINE=22563396; PubMed=12676563;

RX Ruauult M., Ventura M., Galtier N., Brun M.-E., Archidiacono N.,

RA Roizes G., De Sario A.;

RT "BAGE genes generated by juxtapacentromeric reshuffling in the hominidae lineage are under selective pressure.",

RA Genomics 81:391-399(2003).

CC -!- FUNCTION: Unknown. Candidate gene encoding tumor antigens.

CC -!- TISSUE SPECIFICITY: Not expressed in normal tissues except in testis. Expressed in melanoma, bladder and lung carcinomas.

CC -!- MISCELLANEOUS: The ancestral BAGE gene was generated by juxtapacentromeric reshuffling of the MLL3 gene. The BAGE family was expanded by juxtapacentromeric regions of chromosomes 13 and 21 and of exchanges. BAGE family is composed of expressed genes that map to the juxtapacentromeric regions of chromosomes 9, 13, 18 and 21.

- - SIMILARITY: Belongs to the BAGE family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

CC DR EMBL; AF218570; AAL55648.1; - .

CC DR Genew; HGNC:15723; BAGE2.

CC KW Antigen; Multigene family; Signal.

CC SIGNAL 1 17 POTENTIAL.

CC FT CHAIN 18 109 B MELANOMA ANTIGEN 2.

CC SEQUENCE 109 AA; 12114 MW; 51773FAA2C135FF CRC64;

CC SQ

Query Match	83.1%	Score 177;	DB 1;	Length 109;
Best Local Similarity	94.9%	Pred. No. 2	7e-17;	
Matches	37;	Conservative	0;	Mismatches 2;
		Indels	0;	Gaps 0;

Qy 1 MAARAVFIALSAQLQARLMKEESPVWSWRLEPEDGDTAL 39

Db 1 MAAGVFLFLASQQLQARLMKEESPVWSWRLEPEDGDTAL 39

RESULT 4

BGE3_HUMAN	STANDARD;	ERT;	109 AA.
ID BGE3_HUMAN			
AC Q86Y29;			
DT 10-OCT-2003 (Rel. 42, Created)			
DT 10-OCT-2003 (Rel. 42, Last sequence update)			
DT 10-OCT-2003 (Rel. 42, Last annotation update)			
DE B melanoma antigen 3 precursor.			
GN BAGE3.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX [1]			
RN NCBI_TAXID=9606;			
RP SEQUENCE FROM N.A.			
RX MEDLINE:22339465; PubMed=12461691;			
RA Ruault M., Van Der Bruggen P., Brun M.-E., Boyle S., Roizes G.,			
RA Sario A.D., De Sario A.,			
"New BAGE (B melanoma antigen) genes mapping to the juxtapacentromeric regions of human chromosomes 13 and 21 have a cancer/testis profile.";			
RT Eur. J. Hum. Genet. 10:833-840(2002).			
RL [2]			
RP SEQUENCE FROM N.A.			
RX MEDLINE:22563396; PubMed=12576563;			
RA Ruault M., Venturoli M., Galtier N., Brun M.-E., Archidiacono N.,			
RA Roiz G., De Sario A.,			
"BAGE genes generated by juxtapacentromeric reshuffling in the hominidae lineage are under selective pressure.";			
RT Genomics 81:391-399(2003).			
RU -			
CC -I- FUNCTION: Unknown. Candidate gene encoding tumor antigens.			
CC -I- TISSUE_SPECIFICITY: Not expressed in normal tissues except in			
CC testis. Expressed in melanoma, bladder and lung carcinomas.			
CC -I- MISCELLANEOUS: The ancestral BAGE gene was generated by			
CC juxtapacentromeric reshuffling of the MuSL1 gene. The BAGE family was			
CC expanded by juxtapacentromeric movement and/or acrocentric			
CC exchanges. The BAGE family is composed of expressed genes that map to			
CC the juxtapacentromeric regions of chromosomes 13 and 21 and of			
CC unexpressed gene fragments that scattered in the juxtapacentromeric			
CC regions of several chromosomes, including chromosomes 9, 13, 18			
CC and 21.			
CC -I- SIMILARITY: Belongs to the BAGE family.			
CC			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

---

DR EMBL; AF339514; AA032634; 1; -.

DR Gene; HGNC:15728; BAGE3;

KW Antigen; Multigen family; Signal.

FT SIGNAL. 17 POTENTIAL.

FT 18 109 B MELANOMA ANTIGEN 3.

FT CHAIN 12112 MW; A477B7A8FC3A2F4E CRC64;

FT SEQUENCE 109 AA; 3.

Query Match 83.1%; Score 177; DB 1; Length 109;

Best Local Similarity 94.9%; Pred. No. 2.7e-17;

Matches 37; Conservative 0; Mismatches 2; Indels 0; Caps 0; Del 0;

Db 1 MAAGYVFLASAQLICARLKEEESPVSWLEPEDDTAL 39

1 MAARRAFLALSAQQLICARLKEEESPVSWLEPEDDTAL 39

---

RESULT 5

	BGE4_HUMAN	STANDARD	PRT;	39 AA.
ID	BGE4_HUMAN			
AC	Q86Y28;			
DT	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DB	B melanoma antigen 4 precursor.			
GN	BAGE4.			
HS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TAXID	9606;			
[1]				
RN	SEQUENCE FROM N.A.			
RP	SEQUENCE=Testis;			
RC	TISSUE=Testis;			
RA	Ruault M., Rocchi M., Boyle S., Roizes G., van der Bruggen P., De Sario A.;			
RT	"BAGE is a family of centromeric genes coding for tumor antigens.";			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.			
-!-	FUNCTION: Unknown. Candidate gene encoding tumor antigens.			
-!-	TISSUE SPECIFICITY: Not expressed in normal tissues except in testis. Expressed in melanoma, bladder and lung carcinomas.			
-!-	MISCELLANEOUS: The ancestral BAGE gene was generated by juxtapacentromeric reshuffling of the MLL3 Gene. The BAGE family was expanded by juxtapacentromeric movement and/or acrocentric exchanges. BAGE family is composed of expressed genes that map to the juxtapacentromeric regions of chromosomes 13 and 21 and of unexpressed gene fragments that scattered in the juxtapacentromeric regions of several chromosomes, including chromosomes 9, 13, 18 and 21.			
-!-	SIMILARITY: Belongs to the BAGE family.			



cytokine receptor preferentially expressed in hematopoietic cells.";  
RJ Biochem. Biophys. Res. Commun. 272:224-229 (2000).  
[2] SEQUENCE FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPlicing.  
RC TISSUE=lymphocytes;  
RX MEDLINE=20197866; PubMed=110733486;  
RA Fujio K., Nobuka T., Kojima T., Kawashima T., Yahata T., Yamamoto K., Mishimura T.,  
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Yamamoto K., Mishimura T.,  
RA Kitamura T.;  
RT "Molecular cloning of a novel type I cytokine receptor similar to the common gamma chain.";  
RT Blood 95:2204-2211 (2000).  
RL [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=C57BL/6; TISSUE=lymphocytes;  
RX MEDLINE=20432254; PubMed=110974032;  
RA Park L.S., Martin J., Garka K., Gliniak B., Di Santo J.P., Muller W.,  
RA Copeland N.G., Jenkins N.A., Parr A.G.,  
RA Ziegler S.F., Morrissey P.J., Paxton R., Sims J.E.,  
RT "Cloning of the murine thymic stromal lymphopoietin (TSLP) receptor."  
RT Formation of a functional heteromeric complex requires interleukin-7 receptor";  
RT J. Exp. Med. 192:659-670 (2000).  
RL [4]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=21177258; PubMed=1088176;  
RA Pandit A., Ozaki K., Baumann H., Levin S.D., Puel A., Farr A.G.,  
RA Ziegler S.F., Leonard W.J., Lodish H.F.,  
RT "Cloning of a receptor subunit required for signaling by thymic stromal lymphopoietin.";  
RL Nat. Immunol. 1:59-64 (2000).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=FVB/N;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Straubberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuller G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Schuller G.D.,  
RA Altzinger S.F., Zeeberg B., Buston R.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Massina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Scarce M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Yoshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahay S.J.,  
RA Bosack S.A., McEvany P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villazon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,  
RA Schnech A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [6]  
RP SEQUENCE OF 234-359 FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=embryonic stem cells;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamoto I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirali L.M., Stabili F., Suzuki R., Tomita M., Baldarelli R., Barsh G.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,  
RA Norzono P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

Sasaki H., Sato K., Schoenbach C., Suya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wijnhaw-Boris A., Yoshida K., Kawaji H., Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690 (2001).  
CC -!- FUNCTION: Receptor for thymic stromal lymphopoietin (TSLP). Forms a functional complex with TSLP and IL7R which is capable of stimulating cell proliferation through activation of STAT3 and STAT5. Also activates JAK2. Implicated in the development of the hematopoietic system.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1 and isoform 3). Secreted (isoform 2).  
CC -!- ALTERNATIVE PRODUCTS:  
CC -!- Event=Alternative splicing; Named isoforms=3;  
Comment=Additional isoforms seem to exist;  
Name=1; IsoID=Q8CII9-1; Sequence=Displayed;  
Name=2; Synonyms=Soluble CRIM-2;  
Name=3; IsoID=Q8CII9-2; Sequence=VSP\_008788; VSP\_008789;  
Name=4; IsoID=Q8CII9-3; Sequence=VSP\_008790;  
Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: High level of expression in liver, lung and testis. Also expressed in heart, brain, spleen, thymus and bone marrow. Highly expressed in progenitors and myeloid cells. Isoform 2 is expressed in primary hematopoietic cells.  
CC -!- INDUCTION: Upregulated in the myeloid 32D cell line by granulocyte colony stimulating factor (G-CSF).  
CC -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein folding and thereby efficient intracellular transport and cell-surface receptor binding.  
CC -!- DOMAIN: The BOX 1 motif is important for association with JAKs.  
CC -!- SIMILARITY: Belongs to the type I cytokine family of receptors.  
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.  
CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

CC -!- DR EMBL; AB039945; BAA92604; 1;  
CC -!- DR EMBL; AB031333; BAA92159; 1;  
CC -!- DR EMBL; AP232936; AAF81676; 1;  
CC -!- DR EMBL; AF019353; AAF82189; 1;  
CC -!- DR EMBL; BC023738; AAH23788; 1;  
CC -!- DR EMBL; AK010291; BAB26827; 1;  
CC -!- DR MGDB; MGI:188506; TSLP;  
CC -!- DR InterPro; IPB003961; FN\_III.  
CC -!- DR Pfam; PF00041; fn3; 1;  
CC -!- DR SMART; SM00050; FN2; 1;  
CC -!- DR PROSITE; PS01355; HEMATOPO\_REC\_S-FL; FALSE\_NEG.  
CC -!- DR Receptor; Signal; Transmembrane; Glycoprotein; Alternative splicing.  
KW POTENTIAL.  
FT SIGNAL; 1  
FT CHAIN 20 359  
FT DOMAIN 233 253  
FT TRANSM 233 232  
FT DOMAIN 254 359  
FT DOMAIN 117 208  
FT SITE 201 205  
FT SITE 262 269  
FT DISULFID 68 82  
FT CARBOHYD 53 53  
FT CARBOHYD 53 53  
FT CARBOHYD 122 122  
FT VARSPLIC 217 217  
FT VARSPLIC 217 217  
FT VARSPLIC 218 359  
FT VARSPLIC 217 217  
FT VARSPLIC 217 217

FT	CONFLICT	15	15	/FTId=YSP 008790;				
FT	CONFLICT	51	51	A -> T (IN REF. 1).				
FT	CONFLICT	87	87	G -> S (IN REF. 4 AND 5).				
FT	CONFLICT	179	179	A -> G (IN REF. 4).				
FT	CONFLICT	309	309	A -> V (IN REF. 1 AND 3).				
SQ	SEQUENCE	359 AA;	37761 MW;	F9C521C54B4AC9DD CRC64;				
Query Match		26.1%	Score 55.5;	DB 1; Length 359;				
Best Local Similarity		44.0%;	Pred. No. 4.6;					
Matches	11;	Conservative	7;	Mismatches 4;	Indels 3;	Gaps 1;		
Qy	21	KEEPSPVV--SVPRLPEPDGITALCFI 42	:					
Db	301	EEDDLIHPKAKRVEPDPGTSLCTV 325	:					
<b>RESULT 8</b>								
Y166_ARCFU	ID	YL66_ARCFU	STANDARD;	PRT;	154 AA.			
AC	028116;							
DT	16-OCT-2001	(Rel. 40,	Created)					
DT	16-OCT-2001	(Rel. 40,	Last sequence update)					
DT	16-OCT-2001	(Rel. 40,	Last annotation update)					
DE	Hypothetical protein	AF2166.						
GN	OS	Archaeoglobus fulgidus.						
OC	Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;							
OC	Archaeoglobaceae; Archaeoglobus.							
OX	NCBI_TaxID=2234;							
RN	SEQUENCE FROM N.A.							
RC	STRAIN=VC-16 / DSM 4304 / ATCC 49558;							
RX	MEDLINE=98049343; PubMed=9383475;							
RA	Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,							
RA	Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,							
RA	Richardson D.L., Kerlavage A.R., Graham D.E., Kurpides N.C.,							
RA	Fleischmann R.D., Clayton R.A., Kerlavage A.R., Graham D.E., Kurpides N.C.,							
RA	Kirkness E.F., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,							
RA	Peterson J.D., McNeil L.K., Adams M.D., Loftus B., Overbeek R., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,							
RA	Cotton M.D., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,							
RA	Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,							
RA	Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,							
RA	Venter J.C.;							
RT	The complete genome sequence of the hyperthermophilic, sulphate-							
RT	reducing archaeon <i>Archaeoglobus fulgidus</i> .";							
CC	- - SUBCELLULAR LOCATION: Integral membrane protein (Potential).							
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration							
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -							
CC	use by non-profit institutions as long as its content is in no way							
CC	modified and this statement is not removed. Usage by and for commercial							
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@sb-sib.ch).							
DR	EMBL; AE000955; AB89089.1;							
DR	TIGR; AF2166; -							
KW	Hypothetical protein; Transmembrane; Complete proteome.							
FT	TRANSMEM	15	37	POTENTIAL.				
FT	TRANSMEM	58	80	POTENTIAL.				
FT	TRANSMEM	95	116	POTENTIAL.				
SQ	SEQUENCE	123	145	POTENTIAL.				
SEQUENCE	154 AA;	17266 MW;	1CE41361CDC8FA468 CRC64;					
Query Match		25.8%	Score 55;	DB 1; Length 154;				
Best Local Similarity		34.2%;	Pred. No. 2.2;					
Matches	13;	Conservative	10;	Mismatches 15;	Indels 0;	Gaps 0;		
Qy	2 AARVFLLASAQLLQARILKEEESPVVWSRLPEPDGTAL 39							
Db	24 ATNAFLAYALITKSKVDKIKRPPIVTMLLVPLDFAM 61							
Db	24 TNPA.							
RESULT 9								
TNPA_ECOLI	ID	TNPA_ECOLI	STANDARD;	PRT;	994 AA.			
AC	P05504;							
DT	01-AUG-1988	(Rel. 08, Created)						
DT	01-AUG-1988	(Rel. 08, Last sequence update)						
DT	16-OCT-2001	(Rel. 40, Last annotation update)						
DE	Transposase for transposon Tn2501.							
GN	TNPA.							
OS	Escherichia coli.							
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;							
OC	Enterobacteriaceae; Escherichia.							
NCBI_TaxID=562;								
OX								
RN	[1]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=K12;							
RX	Medline=88056496; PubMed=2827105;							
RA	Turner A.X., Grinsted J.;							
RT	"DNA sequence of the new category of class II							
RT	transposon, Tr2501."							
RL	Nucleic Acids Res. 15:10049-10049(1987).							
RN	[2]							
RP	SEQUENCE OF 1-204 FROM N.A.							
RX	Medline=87109049; PubMed=3027041;							
RA	Michiels T., Cornelis G., Ellis K., Grinsted J.,							
RA	RT "Tr2501, a component of the lactose transposon Tr951, is an example							
RT	of a new category of class II transposable elements."							
RL	J. Bacteriol. 169:624-631(1987).							
CC	- - FUNCTION: REQUIRED FOR TRANPOSITION OF TRANSPOSON TN2501.							
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration							
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -							
CC	the European Bioinformatics Institute. There are no restrictions on its							
CC	use by non-profit institutions as long as its content is in no way							
CC	modified and this statement is not removed. Usage by and for commercial							
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@sb-sib.ch).							
CC	or send an email to license@sb-sib.ch).							
CC	CC EMBL; M15197; AA27425; -;							
DR	EMBL; Y00502; QAA68555..1; -;							
DR	PIR; S06302; TOECT5;							
DR	InterPro; IPR02513; Transposase_7.							
DR	Pfam; PF01526; Transposase_7..1.							
KW	Transposable element; Transposition; DNA-binding; DNA recombination.							
SQ	SEQUENCE 994 AA; 112750 MW; 73B0489532AE1DA CRCE4;							
Query Match		25.6%; Score 54.5;	DB 1; Length 994;					
Best Local Similarity		38.1%; Pred. No. 19;						
Matches	16;	Conservative	9;	Mismatches 8;	Indels 9;	Gaps 2;		
Qy	1 MAARAVFLAQLAQIAR-----LMK-----ESPVVSNRLEP 33							
Db	: :   :   :   :   :   :   :   :   :   :   :   :   :							
	172 LADRFLASALTQLEKQHKSALDSILSKSERGEQASRLAVLQP 213							
RESULT 10								
CHO_SALTI	ID	CHO_SALTI	STANDARD;	PRT;	293 AA.			
AC	Q8Z655;							
DT	28-PBB-2003 (Rel. 41, Created)							
DT	28-PBB-2003 (Rel. 41, Last sequence update)							
DT	10-Oct-2003 (Rel. 42, Last annotation update)							
DE	Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (UvrC homolog							
GN	CHO CR STY1804 OR T1188.							
OS	Salmonella typhi.							
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;							
OC	Enterobacteriaceae; Salmonella.							



- Welch R.A., Blattner F.R.; "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.;" RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli;" Nature 409:529-533(2001). [2]

SEQUENCE FROM N.A.

STRAIN=O157:H7 / PMID: 0509953; MEDLINE=115631; PubMed=11228796; RX

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo N., Murata T., Tanaka M., Tobe T., Iida T., Nakami H., Honda T., Sasaki A., Ogassawa C., Yasunaga N., Kuwara S., Shiba T., Hattori M., Shinagawa H.; RA "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12."; RT RNA Res. 8:11-22(2001).

-!- FUNCTION: Incises the DNA at the 3' side of a lesion during nucleotide excision repair. Incises the DNA farther away from the lesion than uvrC. Not able to incise the 5' site of a lesion. When a lesion remains because uvrC is not able to induce the 3', incision, cho incises the DNA. Then uvrC makes the 5' incision. The combined action of cho and uvrC broadens the substrate range of nucleotide excision repair (BY similarity).

-!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRc FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

-!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRc FAMILY.

This SWISS-PROT entry is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

-!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRc FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

SEQUENCE FROM N.A.

STRAIN=K12 / PMID: 17024 (2002)

-!- FUNCTION: Incises the DNA at the 3' side of a lesion during nucleotide excision repair. Incises the DNA farther away from the lesion than uvrC. Not able to incise the 5' site of a lesion. When a lesion remains because uvrC is not able to induce the 3', incision, cho incises the DNA. Then uvrC makes the 5' incision. The combined action of cho and uvrC broadens the substrate range of nucleotide excision repair (BY similarity).

-!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRc FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE016761; AAC80599.1; -.

DR InterPro; IPR000105; UvrC\_N; PFAM; PF0151; Exc endo\_N; 1.

SMART; SM00465; GIYC; 1.

PROSITE; PS00164; UVRc\_1; 1.

SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33694 MW; 9278379AB389130E CRC64;

Query Match 25.4%; Score 54; DB 1; Length 295; Best Local Similarity 50.0%; Pred. No. 6; Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 10 I<sub>1</sub>S<sub>2</sub>A<sub>3</sub>Q<sub>4</sub>L<sub>5</sub>O<sub>6</sub>R<sub>7</sub>M<sub>8</sub>K<sub>9</sub>E<sub>10</sub>P<sub>11</sub>V<sub>12</sub>V<sub>13</sub>W<sub>14</sub>R<sub>15</sub>I<sub>16</sub> 31

Db 89 I<sub>1</sub>G<sub>2</sub>A<sub>3</sub>L<sub>4</sub>L<sub>5</sub>E<sub>6</sub>R<sub>7</sub>K<sub>8</sub>I<sub>9</sub>Q<sub>10</sub>P<sub>11</sub>F<sub>12</sub>N<sub>13</sub>K<sub>14</sub>L<sub>15</sub>R<sub>16</sub> 110

RESULT 14

CHO\_ECOLI STANDARD; PRT; 295 AA.

ID CHO\_ECOLI ID CHO\_ECOLI STANDARD; PRT; 295 AA.

AC P76213; AC P76213;

DR 28-FEB-2003 (Rel. 41, Created) DT 28-FEB-2003 (Rel. 41, Last sequence update)

DR SMART; SM00465; GIYC; 1. DT 10-OCT-2003 (Rel. 42, Last annotation update)

DR PROSITE; PS00164; UVRc\_1; 1. DE Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (UvrC homolog protein).

DR SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome. DE Protein.

DR 33724 MW; 7CBDF68E4AEFF4321 CRC64; GN MAU B.; SHAO Y.; CHO OR B1741.

DR SEQUENCE 295 AA; 33724 MW; 7CBDF68E4AEFF4321 CRC64; GN Escherichia coli.

DR Query Match 25.4%; Score 54; DB 1; Length 295; Best Local Similarity 50.0%; Pred. No. 6; Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 10 I<sub>1</sub>S<sub>2</sub>A<sub>3</sub>Q<sub>4</sub>L<sub>5</sub>O<sub>6</sub>R<sub>7</sub>M<sub>8</sub>K<sub>9</sub>E<sub>10</sub>P<sub>11</sub>V<sub>12</sub>V<sub>13</sub>W<sub>14</sub>R<sub>15</sub>I<sub>16</sub> 31

Db 89 I<sub>1</sub>G<sub>2</sub>A<sub>3</sub>L<sub>4</sub>L<sub>5</sub>E<sub>6</sub>R<sub>7</sub>K<sub>8</sub>I<sub>9</sub>Q<sub>10</sub>P<sub>11</sub>F<sub>12</sub>N<sub>13</sub>K<sub>14</sub>L<sub>15</sub>R<sub>16</sub> 110

RESULT 13

CHO\_ECOL6 STANDARD; PRT; 295 AA.

ID CHO\_ECOL6 ID CHO\_ECOL6 STANDARD; PRT; 295 AA.

AC P59419; AC P59419 / MG1655;

DT 10-OCT-2003 (Rel. 42, Created) DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update) DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (UvrC homolog protein).

GN CHO OR C2140. GN "The complete genome sequence of Escherichia coli K-12.";

OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; OC Enterobacteriaceae; Escherichia. OC NCBI\_TaxID=217992; RN [1]

RP FUNCTION. RN [2]

SEQUENCE FROM N.A.

STRAIN=K12 / CPT073 / ATCC 700928; RX MEDLINE=238834; PubMed=12471157; RX

RA Welch R.A.; Burland V.; Plunkett G. III; Redford P.; Roesch P.; RA Welch R.A.; Burland V.; Plunkett G. III; Redford P.; Roesch P.; RA Riley M.; Collado-Vides J.; Glasner J.D.; Rode C.K.; Mayhew G.F.; RA Gregor J.; Davis N.W.; Kirkpatrick H.A.; Goeden M.A.; Rose D.J.; RA Mau B.; Shao Y.; Cho OR B1741.

RA "The complete genome sequence of Escherichia coli K-12."; RN Science 277:1453-1474(1997).

RP FUNCTION. RX MEDLINE=18119418; PubMed=11818552;

RA "Cho, a second endonuclease involved in Escherichia coli nucleotide excision repair."; RX MEDLINE=9746617; PubMed=9278503;

RA Blattner F.R.; Plunkett G. III; Bloch C.A.; Perna N.T.; Burland V.; RA Riley M.; Collado-Vides J.; Glasner J.D.; Rode C.K.; Mayhew G.F.; RA Gregor J.; Davis N.W.; Kirkpatrick H.A.; Goeden M.A.; Rose D.J.; RA Mau B.; Shao Y.; Cho OR B1741.

RA "Incises the DNA at the 3' side of a lesion during nucleotide excision repair. Incises the DNA farther away from the lesion than uvrC. Not able to incise the 5' site of a lesion. When a lesion remains because uvrC is not able to induce the 3', incision, cho incises the DNA. Then uvrC makes the 5' incision. The combined action of cho and uvrC broadens the substrate range of nucleotide excision repair (BY similarity)."; RX MEDLINE=991457-1472(2002); RX

vitro, the incision activity of cho is uvrA and uvrB dependent. When a lesion remains because uvrC is not able to induce the 3' incision, cho incises the DNA. Then uvrC makes the 5' incision. The combined action of cho and uvrC broadens the substrate range of nucleotide excision repair.

-!- INDUCTION: Repressed by lexa.

-!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AE00269; AAC74811.1; .

DR PIR; E64933; E64933.

DR Ecogene; EG13933; cho.

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33739 MW; DCEFAADD855EP4327 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination; DNA excision; Complete proteome.

SEQUENCE 295 AA; 33705 MW; 875D7DF594BF4472 CRC64;

SQ DR EMBL; AB15171; ANN307.1; ALT\_INIT.

DR EMBL; AB016283; AAP1670.1; .

DR InterPro; IPR000305; Uvrc\_C\_N.

DR Pfam; PF01541; Exc1\_endo\_N\_1.

DR SMART; SM00465; GIYC\_1\_1.

DR PROSITE; PS50164; UVRC\_1\_1.

KW SOS response; Hydrolase; Excision nuclease; DNA repair; DNA recombination

This Page Blank (uspto)